09831000 .050201

JC17 Rec'd PCT/PTO 02 MAY 2001

WDN/SAS:gte 05/02/01 178-59010 48844.doc

EXPRESS MAIL LABEL NO. EL828141124US
DATE OF DEPOSIT: May 1, 2001

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER 178-59010
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. § 371		U.S. APPLICATION NO. (If known, see 37 C.F.R. § 1.5) 09/831000
INTERNATIONAL APPLICATION NO. PCT/US99/26260	INTERNATIONAL FILING DATE November 5, 1999	PRIORITY DATE CLAIMED November 6, 1998

TITLE OF INVENTION
CLONING OF RHESUS MACAQUE RHADINOVIRUS GENOME AND METHODS FOR ITS USE

APPLICANT(S) FOR DO/EO/US
Scott W. Wong, Michael K. Axthelm, and Robert P. Searles

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. § 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. § 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. § 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. § 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. § 371(c)(2))
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ has been transmitted by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. § 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. § 371(c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. § 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. § 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. § 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 C.F.R. §§ 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. §§ 3.28 and 3.31 and the Recordal fee of $40.00 is included.
13. ☒ A **FIRST** preliminary amendment.
    ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
14. ☒ Sequence Listing: Paper copy, 293 pages; computer readable copy on diskette.
15. ☒ Statement in Compliance with 37 C.F.R. § 1.821(f) verifying identity of above copies.
16. ☐ Other items or information:
    ☐ Written Opinion.
    ☐ Preliminary Examination Report.
    ☐ International Search Report.
    ☐ Copies of References Cited.

24197

09831000 .050201

JC08 Rec'd PCT/PTO 02 MAY 2001

WDN/SAS:gte 05/02/01 178-59010 48844.doc

EXPRESS MAIL LABEL NO. EL828141124US
DATE OF DEPOSIT: May 1, 2001

U.S. APPLICATION NO. (If known, see 37 C.F.R. § 1.5) 09/831000	INTERNATIONAL APPLICATION NO. PCT/US99/26260	ATTORNEY'S DOCKET NUMBER 178-59010

17. ☒ The following fees are submitted:

BASIC NATIONAL FEE (37 C.F.R. §§ 1.492(a)(1)-(5)):

Neither International Preliminary Examination fee (37 C.F.R. § 1.482) nor International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1,000.00**

International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$860.00**

International Preliminary Examination fee (37 C.F.R. § 1.482) not paid to USPTO but International Search fee (37 C.F.R. § 1.445(a)(2)) paid to USPTO **$710.00**

International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$690.00**

International Preliminary Examination fee paid to USPTO (37 C.F.R. § 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**

CALCULATIONS (PTO USE ONLY)

				CALCULATIONS	PTO USE ONLY
ENTER APPROPRIATE BASIC FEE AMOUNT =				$ 860.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 C.F.R. § 1.492(e)).				$ 0.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	34 - 20 =	14	x $18.00	$ 252.00	
Independent Claims	2 - 3 =	0	x $80.00	$ 0.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ $270.00	$ 0.00	
TOTAL OF ABOVE CALCULATIONS =				$ 1,112.00	
☒ Reduction of 1/2 for filing by small entity. Small entity status is claimed for this application.				$ 556.00	
SUBTOTAL =				$ 556.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 Months from the earliest claimed priority date (37 C.F.R. §§ 1.492(f)). +				$	
TOTAL NATIONAL FEE =				$ 556.00	
Fee for recording the enclosed assignment (37 C.F.R. § 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. §§ 3.28, 3.31). **$40.00** per property. +				$	
TOTAL FEES ENCLOSED =				$ 556.00	
				REFUND →	$
				CHARGE →	$

a. ☒ A check in the amount of $ 556.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of $ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Director is hereby authorized to charge any additional fees that may be required, or credit any overpayment, to Deposit Account No. 02-4550. A duplicate copy of this sheet is enclosed.

d. ☒ Please return the enclosed postcard to confirm that the items listed above have been received.

NOTE: **Where an appropriate time limit under 37 C.F.R. § 1.494 or § 1.495 has not been met, a petition to revive (37 C.F.R. § 1.137(a) or (b)) must be filed and granted to restore the application to pending status.**

SEND ALL CORRESPONDENCE TO:

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP
One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, OR 97204-2988

William D. Noonan
SIGNATURE

William D. Noonan, M.D.
NAME

30,878
REGISTRATION NUMBER

cc: Docketing

09/831000
531 Rec'd PCT/PTO 02 MAY 2001

WDN/SAS:gte 05/02/01 48850.doc
PATENT

EXPRESS MAIL LABEL NO. EL828141124US
DATE OF DEPOSIT: May 2, 2001
Attorney Reference Number 178-59010
Application Number

#4A
S.A.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Wong et al.

Application No.: To be assigned

Filed: Herewith

For: CLONING OF RHESUS MACAQUE RHADINOVIRUS GENOME AND METHODS FOR ITS USE

Examiner: To be assigned

Date: May 2, 2001

Art Unit: To be assigned

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this paper and the documents referred to as being attached or enclosed herewith are being deposited with the United States Postal Service on May 2, 2001 as Express Mail Label No. EL828141124US in an envelope addressed to: BOX PCT, COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231.

William D. Noonan

William D. Noonan, M.D.
Attorney for Applicant

BOX PCT
COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

PRELIMINARY AMENDMENT

Prior to examination of the above-identified application, please amend the claims as follows:

Please cancel claims 1-34, and insert the following new claims:

35. (New) An isolated virus (RRV) as deposited with ATCC as deposit accession number VR-2601.

A1

36. (New) A purified virus, having a nucleic acid sequence
(a) shown in SEQ ID NO:1 or
(b) a conservative variant thereof.

37. (New) The purified virus of claim 36, wherein the nucleic acid sequence has at least 95% sequence identity to the nucleic acid sequence shown in SEQ ID NO:1.

38. (New) A purified protein encoded by an open reading frame of the virus of claim 36.

cont. A1

39. (New) A purified protein of claim 38, wherein the protein comprises an amino acid sequence selected from the group consisting of:

(a) an amino acid sequence shown in odd numbered sequences of SEQ ID NOS:3-165; and

(b) amino acid sequences that differ from those specified in (a) by one or more conservative amino acid substitutions wherein the function of the protein is preserved.

40. (New) A purified protein with an amino acid sequence that is at least 95% sequence identity to the sequences specified in claim 39.

41. (New) The purified protein of claim 39, wherein the amino acid sequence is selected from odd numbered sequences within the group consisting of SEQ ID NOS:3-19 and 23-165.

42. (New) An isolated nucleic acid molecule encoding a protein according to claim 39.

43. (New) An isolated nucleic acid molecule according to claim 42, wherein the molecule comprises a sequence selected from the group consisting of even numbered sequences of SEQ ID NOS:2-164.

44. (New) The isolated nucleic acid molecule according to claim 43, wherein the molecule comprises a sequence selected from the group consisting of even numbered sequences of SEQ ID NOS:2-18 and 22-164.

cont.
A1

45. (New) A recombinant nucleic acid molecule comprising a promoter sequence operably linked to a nucleic acid molecule according to claim 42.

46. (New) A cell transformed with a recombinant nucleic acid molecule according to claim 42.

47. (New) A non-human mammal purposefully infected with the virus of claim 36.

48. (New) The mammal of claim 47, wherein the mammal is a primate.

49. (New) An oligonucleotide comprising a sequence selected from the group consisting of:

(a) at least 20 contiguous nucleotides of the nucleic acid sequence of the virus of claim 36;

(b) at least 30 contiguous nucleotides of the nucleic acid sequence of the virus of claim 36; and

(c) at least 50 contiguous nucleotides of the nucleic acid sequence of the virus of claim 36.

50. (New) An isolated nucleic acid molecule that encodes the protein of claim 40.

51. (New) An isolated nucleic acid molecule encoding a protein of claim 40.

52. (New) An isolated nucleic acid molecule encoding all proteins encoded by the virus of claim 36, and having a biological activity of an RRV virus.

cont. A1

53. (New) A method for testing the efficacy of a drug in the treatment of a condition associated with the virus of claim 36, the method comprising:

(a) administering the drug to a non-human primate infected with the virus of claim 36; and

(b) observing the primate to determine if the drug prevents or reduces the presentation of one or more symptoms associated with viral infection.

54. (New) The method of claim 53, wherein the primate is immunocompromised.

55. (New) The method of claim 54, wherein the drug is for the treatment of Kaposi's sarcoma and lymphoproliferative disorders.

56. (New) The method of claim 54, wherein the primate is immuno-compromised as a result of infection by Simian Immunodeficiency Virus (SIV).

57. (New) The method of claim 53, wherein the condition associated with infection with the virus is one or more of B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobinulinemia or autoimmune hemolytic anemia.

58. (New) The method of claim 53, wherein the non-human primate is a Rhesus macaque monkey.

cont. A1

59. (New) A method for producing a non-human primate model for testing potential treatments for a condition associated an infection with the virus of claim 36, comprising

(a) administering a treatment to the primate to render the primate immunocompromised; and

(b) infecting the primate with the virus of claim 36.

60. (New) The method of claim 59, wherein the condition is Kaposi's sarcoma and lymphoproliferative disorders.

61. (New) The method of claim 59, wherein the treatment used to render the primate immuno-compromised is infection with SIV.

62. (New) The method of claim 59, wherein the non-human primate is a Rhesus macaque monkey.

63. (New) A method for testing the efficacy of a candidate vaccine against the virus of claim 36, or conditions associated infection with the virus of claim 36, the method comprising:

(a) administering the vaccine to a subject capable of infection with the virus of claim 36;

(b) inoculating the subject with the virus; and

(c) observing the subject to determine if the vaccine prevents or reduces an incidence of viral infection or presentation of one or more conditions associated with the viral infection.

64. (New) The method of claim 63, wherein the subject is a primate.

65. (New) The method of claim 64, wherein the primate is a non-human primate.

09831000 .050201

WDN/SAS:gte 05/02/01 48850.doc
PATENT

EXPRESS MAIL LABEL NO. EL828141124US
DATE OF DEPOSIT: May 2, 2001
Attorney Reference Number 178-59010
Application Number

cont A1

66. (New) The method of claim 63, wherein the primate is immunocompromised.

67. (New) The method of claim 63, wherein the conditions associated with infection include B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobinulinemia or autoimmune hemolytic anemia.

68. (New) The method of claim 65, wherein the non-human primate is a Rhesus macaque monkey.

CONCLUSION

No new matter is added. Entry of this amendment is respectfully requested prior to examination. If any minor matters remain to be addressed prior to examination, the Examiner is invited to call the undersigned at the telephone number listed below.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP

By William D Noonan
William D. Noonan, M.D.
Registration No. 30,878

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

09/831000

JC08 Rec'd PCT/PTO 02 MAY 2001

WDN/SAS:gte 05/02/01 49081
PATENT

EXPRESS MAIL LABEL NO. EL828141124US
DATE OF DEPOSIT: May 2, 2001
Attorney Reference Number 178-59010
Application Number

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Wong et al.

Application No.: To be assigned

Filed: Herewith

For: CLONING OF RHESUS MACAQUE RHADINOVIRUS GENOME AND METHODS FOR ITS USE

Examiner: To be assigned

Date: May 2, 2001

Art Unit: To be assigned

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this paper and the documents referred to as being attached or enclosed herewith are being deposited with the United States Postal Service on May 2, 2001 as Express Mail Label No. EL828141124US in an envelope addressed to: BOX PCT, COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231.

William D Noonan
William D. Noonan, M.D.
Attorney for Applicant

STATEMENT IN COMPLIANCE WITH 37 C.F.R. § 1.821(f)

BOX PCT
COMMISSIONER FOR PATENTS
Washington, DC 20231

Sir:

In compliance with 37 C.F.R. § 1.821(f), the undersigned declares that the nucleotide and/or amino acid sequences presented in the paper copy of the "Sequence Listing" submitted herewith are the same as the sequences contained in the computer-readable form of said "Sequence Listing." No new matter has been added.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP

By William D Noonan
William D. Noonan, M.D.
Registration No. 30,878

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

CLONING OF RHADINOVIRUS GENOME AND METHODS FOR ITS USE

FIELD OF THE INVENTION

The invention relates to the genome of a rhesus macaque rhadinovirus and provides compositions and methods for the production of animal models useful in assessing the efficacy of drugs and vaccines in the treatment and prevention of conditions associated with infection by the virus, such as Kaposi's sarcoma and lymphoproliferative disorders.

BACKGROUND

Converging lines of evidence indicate that Kaposi's sarcoma-associated herpesvirus (KSHV) is the etiological agent responsible for Kaposi's sarcoma (KS) in individuals with and without HIV infection (Chang et al., 1994, *Science* 266:1865-9; Schalling et al., 1995, *Nature Med.* 7:707-8; Moore & Chang, 1995, *N. Engl. J. Med.* 332:1181-5; Whitby et al., 1995, *Lancet* 346:799-802; Ambroziak et al., 1995, *Science* 268:582-3.; Dupin et al., 1995, *Lancet* 345:761-2.; Chuck et al, 1996, *J. Infect. Dis.* 173:248-51; O'Neill et al., 1996, *J. Clin. Pathol.* 49:306-8; Gao et al., 1996, *Nature Med.* 2:925-8; Kedes et al., 1996, *Nature Med.* 2:918-24; Gao et al., 1996, *N. Engl. J. Med.* 335:233-41). In addition to KS, KSHV is also responsible for other acquired immunodeficiency syndrome (AIDS)-related and non-AIDS-related malignancies, such as primary effusion lymphomas (Cesarman et al., 1995, *N. Engl. J. Med.* 332:1186-91; Nador et al., 1996, *Blood* 88:645-56; Otsuki et al, 1996, *Leukemia* 10:1358-62), and multicentric Castleman's disease (MCD), a B cell proliferation disorder associated with overexpression of IL-6 activity (Soulier et al., 1995, *Blood* 86:1276-80; Yoshizaki et al., 1989, *Blood* 74:1360-7).

More recently, KSHV has been proposed to be involved in multiple myeloma, a B cell abnormality of monoclonal origin (Rettig et al., 1997, *Science* 276:1851-4; Said et al., 1997, *Blood* 90:4278-82; Parravicini et al., 1997, *Science* 278:1969-70; Masood et al., 1997, *Science* 278:1970-1; Whitby et al., 1997, *Science* 278:1971-2; Cottoni et al., 1997, *Science* 278:1972; Brousset et al., 1997, *Science* 278:290-4). Understanding how KSHV is involved in these malignancies is important for the generation of therapies against the spectrum of KSHV-associated diseases.

Testing the efficacy of therapeutics and vaccines against any disease, such as KHSV, is greatly facilitated by the availability of an animal model, such as a non-human primate model, because non-human primates are physiologically very similar to humans. Although such models have been developed for the study of HIV infection (for example, U.S. Patent Nos. 5,212,084 and 5,543,131) none has yet been developed for KSHV infection.

Infection of animals with some herpesviruses, namely *Herpesvirus saimiri* and murine herpesvirus type 68, can cause a lymphoproliferative disorder (LPD). However, these animals are not adequate models of KSHV pathogenesis because they lack certain KSHV genes that may

contribute to viral pathogenesis or influence HIV infection, such as Interleukin 6 (IL-6) and macrophage inflammatory protein 1 (MIP-1) (Albrecht et al., 1992, *J. Virol.* 66:5047-58; Virgin et al., 199?, *J. Virol.* 71:5894-904). Thus, so far the establishment of a non-human primate model for KSHV infection has remained elusive.

The present invention addresses this problem, and others, in the development of animal models for a variety of pathological conditions and diseases.

SUMMARY OF THE DISCLOSURE

Rhesus macaques naturally harbor a virus related to KSHV, referred to as RRV, for rhesus rhadinovirus. Genetic analysis of RRV reveals the presence of an IL-6-like gene in a position analogous to that of the KSHV IL-6. The present disclosure also includes information about pathological conditions associated with RRV infection.

The present invention provides the genomic sequence (nucleotide and amino acid) for the RRV genome and its use for developing a non-human primate model for KSHV infection. The invention includes the genome of the newly isolated Rhesus macaque rhadinovirus, RRV isolate 17577 (referred to herein as RRV), but the invention includes variant RRV viruses and related viruses that infect other species. RRV shows some similarity to human Kaposi's sarcoma-associated herpes virus (KSHV, also called HHV8) and possesses genes for both IL-6 and MIP.

The invention encompasses the isolated polynucleotide genome of RRV as shown in SEQ ID NO 1, and the identified ORFs (open reading frames) of this genome (even-numbered SEQ ID NOS 2-164). Also included within the invention are oligonucleotides comprising at least 15, 20, 30, 40, 50, 70, 100 and 150 consecutive nucleotides of the genome sequence as shown in SEQ ID NO 1. Additionally, the invention encompasses various segments of the RRV genome as shown in SEQ ID NO 1, for instance, segments consisting of 999 nucleotides, for example, from nucleotide 1-999, 1000-1999, 2000-2999, 3000-3999, 4000-4999 and so on until the end of the nucleotide sequence. Proteins and parts of proteins encoded within such segments are also covered by the invention.

The invention also includes purified proteins encoded by the RRV genome, the amino acid sequences of which are shown in odd-numbered SEQ ID NOS 3-165. Proteins that have defined degrees of sequence identity with the proteins of SEQ ID NO 1 are also within the scope of the invention. Such proteins may display, for example, at least 50%, 55%, 60%, 70%, 80%, 90%, 95% or even 98% or greater amino acid sequence identity with the native proteins.

The invention further includes nucleic acids encoding the RRV proteins as well as recombinant nucleic acids that include a promoter operably linked to a nucleic acid that encodes an RRV protein.

Additionally included are isolated nucleic acid molecules of various defined lengths that show at least 50%, 60%, 70%, 80%, 90%, 95%, 98% or 100% sequence identity with an RRV ORF sequence, such as the sequence shown in SEQ ID NO 1, or in one of the other sequence listings. The invention also includes isolated nucleic acid molecules of various defined lengths that hybridize with an ORF as shown in SEQ ID NO 1 under wash hybridization conditions of about 70°C and 0.2 x SSC for 1 hour, or alternatively under less stringent conditions of 65°C, 60°C, or 55°C with about 0.2 to 2 x SSC (with, for instance, about 0.1% SDS) for 1 hour.

Also within the invention are cells and virions that contain the nucleic acid molecules as described above.

Additionally the scope of the invention includes the nucleic acid sequences defined by nucleotides 1 to 11031 of SEQ ID NO 1 and nucleotides 21625 to 133719 of SEQ ID NO 1, and ORFs selected from these nucleic acid sequences. The invention also includes isolated nucleic acid molecules of various defined lengths that show at least 50%, 60%, 70%, 80%, 90%, 95% or 98% sequence identity with, an ORF contained within nucleotides 1-11031 or 21625-133719 of the nucleotide sequence as shown in SEQ ID NO 1. Alternatively, the invention includes at least 15, 20, 30, 40, 50, 70, 100 or 150 consecutive nucleotides within nucleotides 1 to 11031 of SEQ ID NO 1 and nucleotides 21625 to 133719 of SEQ ID NO 1, or within ORFs selected from these nucleic acid sequences.

Also included are isolated nucleic acid molecules of various lengths that hybridize under wash conditions of 70°C and about 0.2 x SSC for 1 hour, or alternatively under less stringent conditions of 65°C, 60°C, or 55°C with from about 0.2 to 2 x SSC (with, for instance, about 0.1% SDS) for 1 hour, with an ORF of nucleotides 1-11031 or 21625-133719 of the nucleotide sequence as shown in SEQ ID NO 1.

Recombinant molecules are also encompassed within the bounds of the invention, and include, for instance, a nucleic acid molecule encoding an RRV protein (or fragments or variants thereof) linked to a non-native nucleic acid sequence such as a promoter. The nucleic acid molecule linked to the promoter may be all or part of an ORF encoding an RRV protein, such as any ORF of SEQ ID NO 1, may be one or more fragments of a DNA sequence selected from the DNA sequence defined by nucleotides 1 to 11031 and nucleotides 21625 to 133719 as shown in SEQ ID NO 1, or DNA sequences encoding variants or fragments of proteins encoded by those sequences.

The present invention also relates to the isolation of a virus (RRV) from a rhesus macaque monkey which, when experimentally introduced into immuno-compromised macaques, produces pathological conditions, such as disease signs and symptoms, that parallel those seen in human subjects infected with KSHV, including lymphoproliferative disorders (LPD), lymphadenopathy, splenomegaly, B cell hyperplasia, autoimmune hemolytic anemia, retroperitoneal fibromatosis (a Kaposi's sarcoma-like mesenchymal proliferative disease of body cavities), and

hypergammaglobulinemia, wherein the virus encodes homologues of IL-6 and MIP-1 which are similar to KSHV.

One aspect of the present invention is the isolated virus, RRV, and related species and other variants thereof. In another aspect of the invention, the virus is used to produce a non-human primate model for KSHV infection, or diseases associated with RRV infection; such a model may be produced, for example, by infecting a non-human primate (such as an immunocompromised non-human primate) with RRV. This model may thus be used to evaluate the efficacy of candidate therapeutics and vaccines for KSHV infection treatment and prophylaxis, or other pathological conditions associated with RRV infection. Although it is not required that the primate be first immuno-compromised and then infected with RRV, particular embodiments of the animal model include both infecting the primate with the virus and rendering it immuno-compromised (or equivalently obtaining an already immunocompromised primate).

In another embodiment, the invention provides a method for testing the efficacy of a drug in the treatment of Kaposi's sarcoma and lymphoproliferative disorders or other pathological conditions associated with RRV infection, by administering the drug to an immuno-compromised non-human primate infected with RRV, and then observing the primate to determine if the drug prevents or reduces the presentation of one or more signs, symptoms, laboratory abnormalities, or other pathological conditions associated with infection with the virus. Such conditions include B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobinulinemia, retroperitoneal fibromatosis (a Kaposi's sarcoma-like mesenchymal proliferative disease of body cavities), and autoimmune hemolytic anemia. The efficacy of a vaccine to prevent KSHV infection, or pathological conditions associated with RRV infection, may similarly be assessed by administering the candidate vaccine to the animal and then attempting to infect the animal with RRV. In particular embodiments, the animal to which the candidate vaccine is administered may be an immunocompromised animal. Failure to infect the animal, when control animals not given the candidate vaccine do become infected, indicates that the vaccine conferred protection.

The foregoing and other objects, features, and advantages of the invention will become more apparent from the following detailed description of several examples which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows a phylogenetic comparison of the gammaherpesviruses Epstein-Barr virus (EBV), Alcelaphine herpesvirus (AHV), Murine herpesvirus (MHV), Herpesvirus saimiri (HVS), Kaposi's sarcoma-associated herpesvirus (KSHV), and Rheusus rhadinovirus 17577 (RRV). It shows that among the known sequenced viruses, RRV is the closest relative to KSHV, using an accepted maximum parsimony method of evaluating evolutionary relationships.

FIG. 2 is a table showing the *Bam*HI, *Eco*RI and *Hind*III restriction fragments of the RRV

genome.

FIG. 3 is a schematic map of the ORFs of RRV. Arrow direction represents direction of transcription.

FIG. 4 is a table showing the size, location and description (similarity to other proteins) of the proteins encoded by the ORFs of RRV.

FIG. 5 is a table showing a comparison of corresponding repeats in RRV and KSHV.

FIG. 6 is a table showing the comparison of interferon regulatory elements encoded by RRV and KSHV.

FIG. 7 is a table comparing the ORFs of RRV, KSHV and HVS. The table shows the start and stop points, the strand (+ or -) from which the ORF is transcribed, the size of the ORFs and the percentage similarity of KSHV and HVS when compared with RRV.

FIGS. 8A-8D are graphs showing CD20+ lymphocytes, antibody response and RhKSHV isolation/detection in macaques infected with SIVmac239 and RRV (A)18483 and (B) 18570 and macaques infected with SIVmac239 only (C) 18503 and (D) 18540. A "+" indicates positive for virus culture or viral DNA, as defined by PCR and Southern blot analysis; "-", negative for virus culture or viral DNA.

FIG. 9 shows the result of the PCR analysis of PBLs and LNMCs from each of the macaques (18483, 18503, 18540 and 18570) for RRV DNA and β-globin in (A) graphical form and (B) digital form.

FIG. 10 shows the DNA sequence of the RRV ORF that encodes the IL-6 protein. The corresponding translated polypeptide sequence is shown in standard three letter code below the DNA sequence.

FIG. 11 shows the DNA sequence of the RRV ORF that encodes the MIP protein. The corresponding translated polypeptide sequence is shown in standard three letter code below the DNA sequence.

SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and the code for amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO 1 shows the nucleotide sequence of the RRV genome.

SEQ ID NO 2 shows the cDNA nucleotide sequence of RRV R1, corresponding to nucleotides 1353-2624 of SEQ ID NO 1.

SEQ ID NO 3 shows the amino acid sequence of the RRV R1 protein.

SEQ ID NO 4 shows the cDNA nucleotide sequence of RRV ORF 2, corresponding to the complement of nucleotides 2692-3258 of SEQ ID NO 1, which encodes dihydrofolate reductase,

and which has some similarity to Kaposi's sarcoma-associated herpesvirus (KSHV) ORF 2.

SEQ ID NO 5 shows the amino acid sequence of the ORF2 protein, dihydrofolate reductase protein, which has some similarity to KSHV ORF 2 protein.

SEQ ID NO 6 shows the cDNA nucleotide sequence of RRV ORF 4, corresponding to nucleotides 3676-5613 of SEQ ID NO 1, which encodes complement binding protein, and which has some similarity to KSHV ORF 4.

SEQ ID NO 7 shows the amino acid sequence of the RRV ORF 4 protein, complement binding protein, corresponding to nucleotides 6045-9443 of SEQ ID NO 1, and which has some similarity to KSHV ORF 4 protein.

SEQ ID NO 8 shows the cDNA nucleotide sequence of RRV ORF 6, corresponding to nucleotides 6045-9443 of SEQ ID NO 1, which encodes ssDNA binding protein, and which has some similarity to KSHV ORF 6.

SEQ ID NO 9 shows the amino acid sequence of the RRV ORF 6 protein, ssDNA binding protein, which has some similarity to KSHV ORF 6 protein.

SEQ ID NO 10 shows the cDNA nucleotide sequence of RRV ORF 7, corresponding to nucleotides 9468-11528 of SEQ ID NO 1, which encodes a transport protein, and which has some similarity to KSHV ORF 7.

SEQ ID NO 11 shows the amino acid sequence of the RRV ORF 7 protein, transport protein, which has some similarity to KSHV ORF 7 protein.

SEQ ID NO 12 shows the cDNA nucleotide sequence of RRV ORF 8, corresponding to nucleotides 11515-14004 of SEQ ID NO 1, which encodes glycoprotein B, and which has some similarity to KSHV ORF 8.

SEQ ID NO 13 shows the amino acid sequence of the RRV ORF 8 protein, glycoprotein B protein, which has some similarity to KSHV ORF 8 protein.

SEQ ID NO 14 shows the cDNA nucleotide sequence of RRV ORF 9, DNA polymerase protein, corresponding to nucleotides 14122-17166 of SEQ ID NO 1, which has some similarity to KSHV ORF 9.

SEQ ID NO 15 shows the amino acid sequence of the RRV ORF 9 protein, DNA polymerase protein, which has some similarity to KSHV ORF 9.

SEQ ID NO 16 shows the cDNA nucleotide sequence of RRV ORF 10, corresponding to nucleotides 17261-18511 of SEQ ID NO 1, which has some similarity to KSHV ORF 10.

SEQ ID NO 17 shows the amino acid sequence of the RRV ORF 10 protein, which has some similarity to KSHV ORF 10.

SEQ ID NO 18 shows the cDNA nucleotide sequence of RRV ORF 11, corresponding to nucleotides 18520-19749 of SEQ ID NO 1, which has some similarity to KSHV ORF 11.

SEQ ID NO 19 shows the amino acid sequence of the RRV ORF 11 protein, which has some similarity to KSHV ORF 11.

SEQ ID NO 20 shows the cDNA nucleotide sequence of RRV R2, corresponding to the complement of nucleotides 19921-20544 of SEQ ID NO 1, which has some similarity to the Kaposi's sarcoma-associated IL-6 gene.

SEQ ID NO 21 shows the amino acid sequence of the RRV R2 protein which has some similarity to IL-6.

SEQ ID NO 22 shows the cDNA nucleotide sequence of RRV ORF 70, thymidylate synthase, corresponding to the complement of nucleotides 20777-21778 of SEQ ID NO 1, and which has some similarity to KSHV ORF 70.

SEQ ID NO 23 shows the amino acid sequence of the RRV ORF 70 protein, thymidylate synthase, which has some similarity to KSHV ORF 70 protein.

SEQ ID NO 24 shows the cDNA nucleotide sequence of RRV R3, corresponding to the complement of nucleotides 22245-22592 of SEQ ID NO 1, which has some similarity to the KSHV K4 viral MIP gene.

SEQ ID NO 25 shows the amino acid sequence of the RRV R3 protein, which has some similarity to KSHV K4 viral MIP protein.

SEQ ID NO 26 shows the cDNA nucleotide sequence of RRV ORF 16, a Bcl-2 homolog, corresponding to nucleotides 26846-27409 of SEQ ID NO 1, which has some similarity to KSHV ORF 16.

SEQ ID NO 27 shows the amino acid sequence of the RRV ORF 16 protein, a Bcl-2 protein homolog, which has some similarity to KSHV ORF 16 protein.

SEQ ID NO 28 shows the cDNA nucleotide sequence of RRV ORF 17, corresponding to the complement of nucleotides 27515-29125 of SEQ ID NO 1, encoding a capsid protein, which has some similarity to KSHV ORF 17.

SEQ ID NO 29 shows the amino acid sequence of the RRV ORF 17 protein, a capsid protein, which has some similarity to KSHV ORF 17 protein.

SEQ ID NO 30 shows the cDNA nucleotide sequence of RRV ORF 18, corresponding to nucleotides 28998-29897 of SEQ ID NO 1, which has some similarity to KSHV ORF 18.

SEQ ID NO 31 shows the amino acid sequence of the RRV ORF 18 protein, which has some similarity to KSHV ORF 18 protein.

SEQ ID NO 32 shows the cDNA nucleotide sequence of RRV ORF 19, corresponding to the complement of nucleotides 29905-31548 of SEQ ID NO 1, which encodes a tegument protein, and which has some similarity to KSHV ORF 19.

SEQ ID NO 33 shows the amino acid sequence of the RRV ORF 19 protein, a tegument protein, which has some similarity to KSHV ORF 19 protein.

SEQ ID NO 34 shows the cDNA nucleotide sequence of RRV ORF 20, corresponding to the complement of nucleotides 31043-32095 of SEQ ID NO 1, and which has some similarity to KSHV ORF 20.

SEQ ID NO 35 shows the amino acid sequence of the RRV ORF 19 protein, which has some similarity to KSHV ORF 19 protein.

SEQ ID NO 36 shows the cDNA nucleotide sequence of RRV ORF 21, corresponding to nucleotides 32094-33767 of SEQ ID NO 1, which encodes a thymidine kinase protein, and which has some similarity to KSHV ORF 21.

SEQ ID NO 37 shows the amino acid sequence of the RRV ORF 21 protein, a thymidine kinase protein, which has some similarity to KSHV ORF 21 protein.

SEQ ID NO 38 shows the cDNA nucleotide sequence of RRV ORF 22, corresponding to nucleotides 33754-35868 of SEQ ID NO 1, and which encodes a glycoprotein H protein, and which has some similarity to KSHV ORF 22.

SEQ ID NO 39 shows the amino acid sequence of the RRV ORF 22 protein, a glycoprotein H protein, which has some similarity to KSHV ORF 22 protein.

SEQ ID NO 40 shows the cDNA nucleotide sequence of RRV ORF 23, corresponding to the complement of nucleotides 35865-37073 of SEQ ID NO 1, which has some similarity to KSHV ORF 23.

SEQ ID NO 41 shows the amino acid sequence of the RRV ORF 23 protein, which has some similarity to KSHV ORF 23 protein.

SEQ ID NO 42 shows the cDNA nucleotide sequence of RRV ORF 24, corresponding to the complement of nucleotides 37123-39321 of SEQ ID NO 1, and which has some similarity to KSHV ORF 24.

SEQ ID NO 43 shows the amino acid sequence of the RRV ORF 24 protein, which has some similarity to KSHV ORF 24 protein.

SEQ ID NO 44 shows the cDNA nucleotide sequence of RRV ORF 25, corresponding to nucleotides 39323-43459 of SEQ ID NO 1, which encodes a major capsid protein, and which has some similarity to KSHV ORF 25.

SEQ ID NO 45 shows the amino acid sequence of the RRV ORF 25 protein, a major capsid protein, which has some similarity to KSHV ORF 25 protein.

SEQ ID NO 46 shows the cDNA nucleotide sequence of RRV ORF 26, corresponding to nucleotides 43491-44408 of SEQ ID NO 1, which encodes a capsid protein, and which has some similarity to KSHV ORF 26.

SEQ ID NO 47 shows the amino acid sequence of the RRV ORF 26 protein, a capsid protein, which has some similarity to KSHV ORF 26 protein.

SEQ ID NO 48 shows the cDNA nucleotide sequence of RRV ORF 27, corresponding to nucleotides 44433-45242 of SEQ ID NO 1, and which has some similarity to KSHV ORF 27.

SEQ ID NO 49 shows the amino acid sequence of the RRV ORF 27 protein, which has some similarity to KSHV ORF 27 protein.

SEQ ID NO 50 shows the cDNA nucleotide sequence of RRV ORF 28, corresponding to

nucleotides 45408-45683 of SEQ ID NO 1, and which has some similarity to KSHV ORF 28.

SEQ ID NO 51 shows the amino acid sequence of the RRV ORF 28 protein, which has some similarity to KSHV ORF 28 protein.

SEQ ID NO 52 shows the cDNA nucleotide sequence of RRV ORF 29b, corresponding to the complement of nucleotides 45733-46779 of SEQ ID NO 1, and which has some similarity to KSHV ORF 29b.

SEQ ID NO 53 shows the amino acid sequence of the RRV ORF 29b, which has some similarity to KSHV ORF 29b protein.

SEQ ID NO 54 shows the cDNA nucleotide sequence of RRV ORF 30, corresponding to nucleotides 46905-47135 of SEQ ID NO 1, and which has some similarity to KSHV ORF 30.

SEQ ID NO 55 shows the amino acid sequence of the RRV ORF 30 protein, which has some similarity to KSHV ORF 30 protein.

SEQ ID NO 56 shows the cDNA nucleotide sequence of RRV ORF 31, corresponding to nucleotides 47093-47746 of SEQ ID NO 1, and which has some similarity to KSHV ORF 31.

SEQ ID NO 57 shows the amino acid sequence of the RRV ORF 31, protein which has some similarity to KSHV ORF 31 protein.

SEQ ID NO 58 shows the cDNA nucleotide sequence of RRV ORF 32, corresponding to nucleotides 47683-49077 of SEQ ID NO 1, and which has some similarity to KSHV ORF 32.

SEQ ID NO 59 shows the amino acid sequence of the RRV ORF 32 protein, which has some similarity to KSHV ORF 32 protein.

SEQ ID NO 60 shows the cDNA nucleotide sequence of RRV ORF 33, corresponding to nucleotides 49049-50059 of SEQ ID NO 1, and which has some similarity to KSHV ORF 33.

SEQ ID NO 61 shows the amino acid sequence of the RRV ORF 33 protein, which has some similarity to KSHV ORF 33 protein.

SEQ ID NO 62 shows the cDNA nucleotide sequence of RRV ORF 29a, corresponding to the complement of nucleotides 49977-50960 of SEQ ID NO 1, and which has some similarity to KSHV ORF 29a.

SEQ ID NO 63 shows the amino acid sequence of the RRV ORF 29a protein, which has some similarity to KSHV ORF 29a protein.

SEQ ID NO 64 shows the cDNA nucleotide sequence of RRV ORF 34, corresponding to nucleotides 50959-51942 of SEQ ID NO 1, and which has some similarity to KSHV ORF 34.

SEQ ID NO 65 shows the amino acid sequence of the RRV ORF 34 protein, which has some similarity to KSHV ORF 34 protein.

SEQ ID NO 66 shows the cDNA nucleotide sequence of RRV ORF 35, corresponding to nucleotides 51923-52372 of SEQ ID NO 1, and which has some similarity to KSHV ORF 35.

SEQ ID NO 67 shows the amino acid sequence of the RRV ORF 35 protein, which has some similarity to KSHV ORF 35 protein.

SEQ ID NO 68 shows the cDNA nucleotide sequence of RRV ORF 36, corresponding to nucleotides 52278-53585 of SEQ ID NO 1, which encodes a kinase, and which has some similarity to KSHV ORF 36.

SEQ ID NO 69 shows the amino acid sequence of the RRV ORF 36 protein, a kinase, which has some similarity to KSHV ORF 36 protein.

SEQ ID NO 70 shows the cDNA nucleotide sequence of RRV ORF 37, corresponding to nucleotides 53566-55008 of SEQ ID NO 1, which encodes an alkaline exonuclease, and which has some similarity to KSHV ORF 37.

SEQ ID NO 71 shows the amino acid sequence of the RRV ORF 37 protein, an alkaline exonuclease protein, which has some similarity to KSHV ORF 37 protein.

SEQ ID NO 72 shows the cDNA nucleotide sequence of RRV ORF 38, corresponding to nucleotides 54963-55172 of SEQ ID NO 1, and which has some similarity to KSHV ORF 38.

SEQ ID NO 73 shows the amino acid sequence of the RRV ORF 38 protein, which has some similarity to KSHV ORF 38 protein.

SEQ ID NO 74 shows the cDNA nucleotide sequence of RRV ORF 39, corresponding to the complement of nucleotides 55255-56391 of SEQ ID NO 1, which encodes glycoprotein M, and which has some similarity to KSHV ORF 39.

SEQ ID NO 75 shows the amino acid sequence of the RRV ORF 39 protein, glycoprotein M, which has some similarity to KSHV ORF 39 protein.

SEQ ID NO 76 shows the cDNA nucleotide sequence of RRV ORF 40, corresponding to nucleotides 56526-57932 of SEQ ID NO 1, which encodes helicase/primase, and which has some similarity to KSHV ORF 40.

SEQ ID NO 77 shows the amino acid sequence of the RRV ORF 40 protein, helicase/primase, which has some similarity to KSHV ORF 40 protein.

SEQ ID NO 78 shows the cDNA nucleotide sequence of RRV ORF 41, corresponding to nucleotides 57917-58528 of SEQ ID NO 1, which encodes helicase/primase, and which has some similarity to KSHV ORF 41.

SEQ ID NO 79 shows the amino acid sequence of the RRV ORF 41 protein, helicase/primase, which has some similarity to KSHV ORF 41 protein.

SEQ ID NO 80 shows the cDNA nucleotide sequence of RRV ORF 42, corresponding to the complement of nucleotides 58525-59343 of SEQ ID NO 1, which has some similarity to KSHV ORF 42.

SEQ ID NO 81 shows the amino acid sequence of the RRV ORF 42 protein, which has some similarity to KSHV ORF 42 protein.

SEQ ID NO 82 shows the cDNA nucleotide sequence of RRV ORF 43, corresponding to the complement of nucleotides 59297-61027 of SEQ ID NO 1, which encodes a capsid protein, and which has some similarity to KSHV ORF 43.

SEQ ID NO 83 shows the amino acid sequence of the RRV ORF 43 protein, a capsid protein, which has some similarity to KSHV ORF 43 protein.

SEQ ID NO 84 shows the cDNA nucleotide sequence of RRV ORF 44, corresponding to nucleotides 60966-63338 of SEQ ID NO 1, which encodes helicase/primase, and which has some similarity to KSHV ORF 44.

SEQ ID NO 85 shows the amino acid sequence of the RRV ORF 44 protein, helicase/primase, which has some similarity to KSHV ORF 44 protein.

SEQ ID NO 86 shows the cDNA nucleotide sequence of RRV ORF 45, corresponding to the complement of nucleotides 63379-64437 of SEQ ID NO 1, and which has some similarity to KSHV ORF 45.

SEQ ID NO 87 shows the amino acid sequence of the RRV ORF 45 protein, which has some similarity to KSHV ORF 45 protein.

SEQ ID NO 88 shows the cDNA nucleotide sequence of RRV ORF 46, corresponding to the complement of nucleotides 64479-65246 of SEQ ID NO 1, which encodes uracil DNA glucosidase, and which has some similarity to KSHV ORF 46.

SEQ ID NO 89 shows the amino acid sequence of the RRV ORF 46 protein, uracil DNA glucosidase protein, which has some similarity to KSHV ORF 46 protein.

SEQ ID NO 90 shows the cDNA nucleotide sequence of RRV ORF 47, corresponding to the complement of nucleotides 65222-65731 of SEQ ID NO 1, which encodes glycoprotein L, which has some similarity to KSHV ORF 47.

SEQ ID NO 91 shows the amino acid sequence of the RRV ORF 47 protein, glycoprotein L, which has some similarity to KSHV ORF 47 protein.

SEQ ID NO 92 shows the cDNA nucleotide sequence of RRV ORF 48, corresponding to the complement of nucleotides 65999-67168 of SEQ ID NO 1, and which has some similarity to KSHV ORF 48.

SEQ ID NO 93 shows the amino acid sequence of the RRV ORF 48 protein, which has some similarity to KSHV ORF 48 protein.

SEQ ID NO 94 shows the cDNA nucleotide sequence of RRV ORF 49, corresponding to the complement of nucleotides 67398-68303 of SEQ ID NO 1, and which has some similarity to KSHV ORF 49.

SEQ ID NO 95 shows the amino acid sequence of the RRV ORF 49 protein, which has some similarity to KSHV ORF 49 protein.

SEQ ID NO 96 shows the cDNA nucleotide sequence of RRV ORF 50, corresponding to nucleotides 68494-70038 of SEQ ID NO 1, which encodes a transactivator, and which has some similarity to KSHV ORF 50.

SEQ ID NO 97 shows the amino acid sequence of the RRV ORF 50 protein, a transactivator protein, which has some similarity to KSHV ORF 50 protein.

SEQ ID NO 98 shows the cDNA nucleotide sequence of RRV R4, corresponding to nucleotides 70355-70888 of SEQ ID NO 1.

SEQ ID NO 99 shows the amino acid sequence of the RRV R4 protein.

SEQ ID NO 100 shows the cDNA nucleotide sequence of RRV R5, corresponding to nucleotides 71468-72160 of SEQ ID NO 1.

SEQ ID NO 101 shows the amino acid sequence of the RRV R5 protein.

SEQ ID NO 102 shows the cDNA nucleotide sequence of RRV ORF 52, corresponding to the complement of nucleotides 72401-72820 of SEQ ID NO 1, and which has some similarity to KSHV ORF 52.

SEQ ID NO 103 shows the amino acid sequence of the RRV ORF 52 protein, which has some similarity to KSHV ORF 52 protein.

SEQ ID NO 104 shows the cDNA nucleotide sequence of RRV ORF 53, corresponding to the complement of nucleotides 72884-73198 of SEQ ID NO 1, and which has some similarity to KSHV ORF 53.

SEQ ID NO 105 shows the amino acid sequence of the RRV ORF 53 protein, which has some similarity to KSHV ORF 53 protein.

SEQ ID NO 106 shows the cDNA nucleotide sequence of RRV ORF 54, corresponding to nucleotides 73274-74146 of SEQ ID NO 1, which encodes a dUTPase, and which has some similarity to KSHV ORF 54.

SEQ ID NO 107 shows the amino acid sequence of the RRV ORF 54 protein, a dUTPase protein, which has some similarity to KSHV ORF 54 protein.

SEQ ID NO 108 shows the cDNA nucleotide sequence of RRV ORF 55, corresponding to the complement of nucleotides 74207-74839 of SEQ ID NO 1, and which has some similarity to KSHV ORF 55.

SEQ ID NO 109 shows the amino acid sequence of the RRV ORF 55 protein, which has some similarity to KSHV ORF 55 protein.

SEQ ID NO 110 shows the cDNA nucleotide sequence of RRV ORF 56, corresponding to nucleotides 74851-77337 of SEQ ID NO 1, which encodes a DNA replication protein, and which has some similarity to KSHV ORF 56.

SEQ ID NO 111 shows the amino acid sequence of the RRV ORF 56 protein, a DNA replication protein, which has some similarity to KSHV ORF 56 protein.

SEQ ID NO 112 shows the cDNA nucleotide sequence of RRV ORF 57, corresponding to nucleotides 77578-78906 of SEQ ID NO 1, which encodes an immediate-early gene product, and which has some similarity to KSHV ORF 57.

SEQ ID NO 113 shows the amino acid sequence of the RRV ORF 57 protein, a immediate-early gene product protein, which has some similarity to KSHV ORF 57 protein.

SEQ ID NO 114 shows the cDNA nucleotide sequence of RRV R6, corresponding to the

complement of nucleotides 79266-80513 of SEQ ID NO 1, which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 115 shows the amino acid sequence of the RRV R6 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 116 shows the cDNA nucleotide sequence of RRV R7, corresponding to the complement of nucleotides 80686-81933 of SEQ ID NO 1,which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 117 shows the amino acid sequence of the RRV R7 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 118 shows the cDNA nucleotide sequence of RRV R8, corresponding to the complement of nucleotides 82262-83317 of SEQ ID NO 1, which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 119 shows the amino acid sequence of the RRV R8 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 120 shows the cDNA nucleotide sequence of RRV R9, corresponding to the complement of nucleotides 83491-84252 of SEQ ID NO 1, which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 121 shows the amino acid sequence of the RRV R9 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 122 shows the cDNA nucleotide sequence of RRV R10, corresponding to the complement of nucleotides 85052-86209 of SEQ ID NO 1,which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 123 shows the amino acid sequence of the RRV R10 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 124 shows the cDNA nucleotide sequence of RRV R11, corresponding to the complement of nucleotides 86355-87527 of SEQ ID NO 1,which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 125 shows the amino acid sequence of the RRV R11 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 126 shows the cDNA nucleotide sequence of RRV R12, corresponding to the complement of nucleotides 87894-88961 of SEQ ID NO 1, which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 127 shows the amino acid sequence of the RRV R12 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 128 shows the cDNA nucleotide sequence of RRV R13, corresponding to the complement of nucleotides 89122-90216 of SEQ ID NO 1,which has some similarity to KSHV vIRF K9 gene.

SEQ ID NO 129 shows the amino acid sequence of the RRV R13 protein, which has some similarity to KSHV vIRF K9 protein.

SEQ ID NO 130 shows the cDNA nucleotide sequence of RRV ORF 58, corresponding to the complement of nucleotides 90462-91544 of SEQ ID NO 1,which has some similarity to KSHV ORF 58.

SEQ ID NO 131 shows the amino acid sequence of the RRV ORF 58 protein, which has some similarity to KSHV ORF 58 protein.

SEQ ID NO 132 shows the cDNA nucleotide sequence of RRV ORF 59, corresponding to the complement of nucleotides 91555-92739 of SEQ ID NO 1, which encodes a DNA replication protein, and which has some similarity to KSHV ORF 59.

SEQ ID NO 133 shows the amino acid sequence of the RRV ORF 59 protein, a DNA replication protein, which has some similarity to KSHV ORF 59 protein.

SEQ ID NO 134 shows the cDNA nucleotide sequence of RRV ORF 60, corresponding to the complement of nucleotides 92868-93812 of SEQ ID NO 1, which encodes a small ribonucleotide reductase protein, and which has some similarity to KSHV ORF 60.

SEQ ID NO 135 shows the amino acid sequence of the RRV ORF 60 protein, a small ribonucleotide reductase protein, which has some similarity to KSHV ORF 60 protein.

SEQ ID NO 136 shows the cDNA nucleotide sequence of RRV ORF 61, corresponding to the complement of nucleotides 93794-96160 of SEQ ID NO 1, which encodes a large ribonucleotide reductase protein, and which has some similarity to KSHV ORF 61.

SEQ ID NO 137 shows the amino acid sequence of the RRV ORF 61 protein, a large ribonucleotide reductase protein, which has some similarity to KSHV ORF 61 protein.

SEQ ID NO 138 shows the cDNA nucleotide sequence of RRV ORF 62, corresponding to the complement of nucleotides 96163-97158 of SEQ ID NO 1, which encodes a assembly/DNA maturation protein, and which has some similarity to KSHV ORF 62.

SEQ ID NO 139 shows the amino acid sequence of the RRV ORF 62 protein, a assembly/DNA maturation protein, which has some similarity to KSHV ORF 62 protein.

SEQ ID NO 140 shows the cDNA nucleotide sequence of RRV ORF 63, corresponding to nucleotides 97157-99976 of SEQ ID NO 1, which encodes a tegument protein, and which has some similarity to KSHV ORF 63.

SEQ ID NO 141 shows the amino acid sequence of the RRV ORF 63 protein, a tegument protein, which has some similarity to KSHV ORF 63 protein.

SEQ ID NO 142 shows the cDNA nucleotide sequence of RRV ORF 64, corresponding to nucleotides 99980-107626 of SEQ ID NO 1, which encodes a tegument protein, and which has some similarity to KSHV ORF 64.

SEQ ID NO 143 shows the amino acid sequence of the RRV ORF 64 protein, a tegument protein, which has some similarity to KSHV ORF 64 protein.

SEQ ID NO 144 shows the cDNA nucleotide sequence of RRV ORF 65, corresponding to the complement of nucleotides 107637-108146 of SEQ ID NO 1, which encodes a capsid protein, and which has some similarity to KSHV ORF 65.

SEQ ID NO 145 shows the amino acid sequence of the RRV ORF 65 protein, a capsid protein, which has some similarity to KSHV ORF 65 protein.

SEQ ID NO 146 shows the cDNA nucleotide sequence of RRV ORF 66, corresponding to the complement of nucleotides 108152-109498 of SEQ ID NO 1, which has some similarity to KSHV ORF 66.

SEQ ID NO 147 shows the amino acid sequence of the RRV ORF 66 protein, which has some similarity to KSHV ORF 66 protein.

SEQ ID NO 148 shows the cDNA nucleotide sequence of RRV ORF 67, corresponding to the complement of nucleotides 109524-110198 of SEQ ID NO 1, which encodes a tegument protein, and which has some similarity to KSHV ORF 67.

SEQ ID NO 149 shows the amino acid sequence of the RRV ORF 67 protein, a tegument protein, which has some similarity to KSHV ORF 67 protein.

SEQ ID NO 150 shows the cDNA nucleotide sequence of RRV ORF 68, corresponding to nucleotides 110609-111982 of SEQ ID NO 1, which encodes a glycoprotein, and which has some similarity to KSHV ORF 68.

SEQ ID NO 151 shows the amino acid sequence of the RRV ORF 68 protein, a glycoprotein, which has some similarity to KSHV ORF 68 protein.

SEQ ID NO 152 shows the cDNA nucleotide sequence of RRV ORF 69, corresponding to nucleotides 112004-112897 of SEQ ID NO 1, which has some similarity to KSHV ORF 69.

SEQ ID NO 153 shows the amino acid sequence of the RRV ORF 69 protein, which has some similarity to KSHV ORF 69 protein.

SEQ ID NO 154 shows the cDNA nucleotide sequence of RRV ORF 71, corresponding to the complement of nucleotides 119211-119735 of SEQ ID NO 1, which encodes a FLIP protein, and which has some similarity to KSHV ORF 71.

SEQ ID NO 155 shows the amino acid sequence of the RRV ORF 71 protein, a FLIP protein, which has some similarity to KSHV ORF 71 protein.

SEQ ID NO 156 shows the cDNA nucleotide sequence of RRV ORF 72, corresponding to the complement of nucleotides 119794-120558 of SEQ ID NO 1, which encodes a cyclin D homolog, and which has some similarity to KSHV ORF 72.

SEQ ID NO 157 shows the amino acid sequence of the RRV ORF 72 protein, a cyclin D homolog protein, which has some similarity to KSHV ORF 72 protein.

SEQ ID NO 158 shows the cDNA nucleotide sequence of RRV ORF 73, corresponding to the complement of nucleotides 120866-122212 of SEQ ID NO 1, which encodes a latent nuclear antigen, and which has some similarity to KSHV ORF 73.

SEQ ID NO 159 shows the amino acid sequence of the RRV ORF 73 protein, a latent nuclear antigen, which has some similarity to KSHV ORF 73 protein.

SEQ ID NO 160 shows the cDNA nucleotide sequence of RRV R15, corresponding to nucleotides 122866-123627 of SEQ ID NO 1, which has some similarity to KSHV K14 and ox-2.

SEQ ID NO 161 shows the amino acid sequence of the RRV R15 protein, which has some similarity to KSHV K14 and ox-2.

SEQ ID NO 162 shows the cDNA nucleotide sequence of RRV ORF 74, corresponding to nucleotides 123924-124952 of SEQ ID NO 1, which encodes a G protein coupled receptor, and which has some similarity to KSHV ORF 74.

SEQ ID NO 163 shows the amino acid sequence of the RRV ORF 74 protein, a G protein coupled receptor protein, which has some similarity to KSHV ORF 74 protein.

SEQ ID NO 164 shows the cDNA nucleotide sequence of RRV ORF 75, corresponding to the complement of nucleotides 125057-128953 of SEQ ID NO 1, which encodes a tegument protein, FGARAT, and which has some similarity to KSHV ORF 75.

SEQ ID NO 165 shows the amino acid sequence of the RRV ORF 75 protein, a tegument protein, FGARAT, which has some similarity to KSHV ORF 75 protein.

The cDNA sequences given in each of the even numbered sequences SEQ ID NOs 2-164 are the open reading frames of the RRV, with the nucleotide references in each of those sequences being given with reference to the nucleotide numbers set forth in SEQ ID NO 1.

SEQ ID NOs 166-172 are PCR primers used in the present invention.

SEQ ID NO 173 shows the coding sequence similar to that for MIP without AUG. Nucleic acid numbers correspond to those given in SEQ ID NO 1.

SEQ ID NOs 174-179 show the repeat regions of the RRV genome. Nucleic acid numbers correspond to those given in SEQ ID NO 1.

SEQ ID NOs 180-181 are probes specific for the KSHV thymidylate synthase (TS) gene used for Southern blot hybridization.

SEQ ID NOs 182-183 are oligonucleotide PCR primers specific for the RRV MIP gene.

ATCC DEPOSIT

A Budapest Treaty deposit of RRV 17577 was made with the American Type Culture Collection (ATCC), Manassas, Virginia, on March 12, 1998, and has been accorded ATCC Accession No. VR-2601.

DETAILED DESCRIPTION OF SEVERAL EMBODIMENTS

Abbreviations and Definitions

Animal: Living multicellular vertebrate organisms, a category which includes, for example, humans, non-human primates, mammals, and birds.

Cell: A plant, animal, insect, bacterial, or fungal cell.

Homologs: two nucleotide or amino acid sequences that share a common ancestral sequence and diverged when a species carrying that ancestral sequence split into two species. Homologs frequently show a substantial degree of sequence identity.

IL-6: Interleukin 6. IL-6 is a cytokine known to have pleiotropic immunological effects including anti-inflammatory and immunosuppressive effects (*Human Cytokines*, 1991, pg. 142-167, Blackwell Scientific Publications, Aggarwal and Gutterman, eds). Because IL-6 is a pleiotropic cytokine, IL-6 activity may be measured using a number of bioassays, including stimulation of immunoglobulin production in SKW6-CL4 cells as described by Hirano et al. (*Nature* 324:73-6, 1986) and stimulation of hybridoma cell growth as described by Matsuda et al., 1988 *Eur. J. Immunol.* 18:951-956, both of which are incorporated by reference. As used herein, the term "IL-6 biological activity" refers to the ability of a protein to show activity in at least one of these assay systems

Immuno-compromised: Lacking a normal immune response. Immuno-compromised refers to a condition in which some or all of an animal's immune system is inoperative, leaving the animal with an increased susceptibility to infection or disease. An animal may be rendered immuno-compromised by a biological agent such as, in the case of non-human primates, Simian Immunodeficiency Virus (SIV). Many strains of SIV have been isolated and characterized; any SIV strain that produces an immuno-compromised state can be used in the present invention including, but not limited to, for example, SIVmac239 (Kestler et al., 1990, *Science* 248: 1109-12), SIVmac251 (Daniels et al., 1985, *Science* 228: 1201-4), SIVdeltaB670 (Murphy-Corb et al., 1986, *Nature* 321:435) and SIVmne (Benveniste et al., 1988, *J. Virol.* 62:2091-101). In addition, hybrid SIV/HIV chimeras as known in the field can be employed, as can HIV-2. Simian type D retroviruses (SRVs) which cause an AIDS-like disease in rhesus monkeys, can alternatively be used to immuno-compromise the animals in place of SIV. These viral agents are administered to the animal using conventional means, such as intravenous or intramuscular injection, or oral, intrarectal or intravaginal inoculation (also see Example 24). Either intact viral particles or viral DNA may be administered. As known in the field, plasmid constructs containing the entire SIV genome are infectious when inoculated into animals and so may be employed in place of purified viral DNA.

Alternatively, an animal may be rendered immuno-compromised by administration of agents that target the immune system, including byt not limited to anti-CD3 antibody (CD3 being the T-cell receptor) either alone or conjugated with a toxic moiety, or immunosuppressive compounds including prednisone, azathioprine, cyclosporine A, and cyclophosphamide. Where an immunosuppressive compound such as cyclosporine is employed, an allogenic stimulus (such as a blood transfusion) may be administered with the subsequent administration of RRV to activate infection.

Alternatively, other methods of rendering an animal immuno-compromised may be used, including radiation treatment and surgical intervention.

Isolated: An "isolated" biological component (such as a nucleic acid, peptide or protein) has been substantially separated, produced apart from, or purified away from other biological components in the cell of the organism in which the component naturally occurs, i.e., other chromosomal and extrachromosomal DNA and RNA, and proteins. Nucleic acids, peptides and proteins which have been "isolated" thus include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids, peptides and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

KSHV: Kaposi's sarcoma-associated herpesvirus. KSHV is a herpesvirus associated with (and thought to be the etiological agent of) Kaposi's sarcoma in humans.

Lymphoproliferative Disorder: a group of disorders characterized by proliferation of lymphoid tissue, such as lymphocytic leukemia and malignant lymphoma, and characterized by such signs as lymphocytosis, lymphadenopathy, and splenomegaly.

MIP: macrophage inflammatory protein. The acronym MIP is used to describe a family of cytokines that includes MIP1 (Davatelis et al., 1989, *Science* 243: 1066-8) and MIP2 (U.S. Patent No. 5,145,676). MIPs mediate pleiotropic immunological effects including activation of neutrophils to undergo an oxidative burst. MIPs are also intrinsically pyrogenic. MIP biological activity can be detected and quantified using bioassays as described in Kedal et al. (*Science* 277:1656-9, 1997) and Boshoff et al. (*Science* 278:290-4, 1997) that measure MIP concentrations using HIV inhibition and calcium mobilization, respectively. As used herein, the term "MIP biological activity" refers to the ability of a protein to show activity in at least one of these assay systems.

Non-human primate: Simian primates including chimpanzees, orangutans, baboons, and macaques. Any non-human primate may be used to produce a KSHV-disease animal model by the methods disclosed herein. Thus, in addition to the rhesus macaque models described in detail below, pigtail and cynomologus macaques and baboons may also be used to produce KSHV-disease animal models by the methods disclosed herein.

Oligonucleotide: A linear polynucleotide sequence of up to about 200 nucleotide bases in length, for example a polynucleotide (such as DNA or RNA) which is at least 6 nucleotides, for example at least 15, 25, 50, 100 or even 200 nucleotides long.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in the same reading frame.

ORF: open reading frame. Contains a series of nucleotide triplets (codons) coding for amino acids without any termination codons. These sequences are usually translatable into protein.

PCR: polymerase chain reaction. Describes a technique in which cycles of denaturation, annealing with primer, and then extension with DNA polymerase are used to amplify the number of copies of a target DNA sequence.

Pharmaceutically acceptable carriers: The pharmaceutically acceptable carriers useful in this invention include conventional carriers. Remington's Pharmaceutical Sciences, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes compositions and formulations suitable for pharmaceutical delivery of the viruses, nucleic acids and/or proteins herein disclosed.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol, ethanol, combinations thereof, or the like, as a vehicle. The carrier and composition can be sterile, and the formulation suits the mode of administration. For solid compositions (e.g., powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, sodium saccharine, cellulose, magnesium carbonate, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

The composition can be a liquid solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides.

Probes and primers: Nucleic acid probes and primers may readily be prepared based on the amino acid sequences provided by this invention. A probe is an isolated nucleic acid attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, e.g., in Sambrook et al., *in* Molecular Cloning: A Laboratory Manual, Cold Spring (1989) and Ausubel et al., in Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley-Intersciences (1987).

Primers are short nucleic acids, such as DNA oligonucleotides 10 nucleotides or more in length. Primers may be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other

nucleic-acid amplification methods known in the art.

Methods for preparing and using probes and primers are described, for example, in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor, New York. 1989); Ausubel et al. (Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences. 1987) and Innis et al. (PCR Protocols, A Guide to Methods and Applications, 1990, Innis et al. (eds.), 21-27, Academic Press, Inc., San Diego, California). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of the RRV genome sequence (SEQ ID NO 1). One of skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 20 consecutive nucleotides will anneal to a target with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers may be selected that comprise 20, 25, 30, 35, 40, 50 or more consecutive nucleotides. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 30, 40, 50, 60, 70, 80, 90, 100, or 150 consecutive nucleotides of the disclosed nucleic acid sequences.

Alternatively, such probes and primers may comprise at least 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, or 150 consecutive nucleotides that share a defined level of sequence identity with the disclosed RRV sequence, for instance, at least a 50%, 60%, 70%, 80%, 90%, 95% or 98% sequence identity. Alternatively, such probes and primers may be nucleotide molecules which hybridize under wash conditions of 70°C and about 0.2 x SSC for 1 hour, or alternatively under less stringent conditions of 65°C, 60°C, or 55°C with from about 0.2 to 2 x SSC (with, for instance, about 0.1% SDS) for 1 hour with a portion of the RRV sequence.

Purified: The term purified does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified peptide preparation is one in which the peptide or protein is more enriched than the peptide or protein is in its natural environment within a cell. Preferably, a preparation is purified such that the protein or peptide represents at least 50% of the total peptide or protein content of the preparation.

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques.

RRV 17577: Rhesus macaque rhadinovirus RRV isolate 17577. A Budapest Treaty deposit of RRV 17577 was made with the American Type Culture Collection, Manassas, Virginia,

on March 12, 1998, and has been accorded ATCC Accession No. VR-2601. This virus may be grown on primary rhesus fibroblasts, as described below (see Examples 1 and 14), using standard virological techniques. Alternatively, it may be grown on commercially available rhesus cell lines, including those available from ATCC, such as ATCC CRL-6306 and ATCC CL-160. Infection of a non-human primate with RRV 17577 may be accomplished using any standard method, including intravenous injection (see Examples 13, 23 and 24). Typically, infection is achieved by intravenous injection of around 10^6 plaque forming units (PFUs) of RRV 17577.

RRV: A virus having the virological and immunological characteristics of RRV 17577, and which causes Kaposi's sarcoma in immunocompromised Rheusus monkeys which are infected with the virus. In particular examples, the RRV has at least 85% (for example at least 90%, 95% or 98%) sequence identity to SEQ ID NO 1.

Sequence Identity: The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Homologs or orthologs of the RRV proteins and the corresponding DNA sequences, will possess a relatively high degree of sequence identity when aligned using standard methods. This homology will be more significant when the orthologous proteins or DNAs are derived from species which are more closely related (e.g., human and chimpanzee sequences), compared to species more distantly related (e.g., human and *C. elegans* sequences).

Typically, RRV orthologs are at least 50% identical at the nucleotide level and at least 50% identical at the amino acid level when comparing RRV to an orthologous RRV sequences.

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins & Sharp, *Gene*, 73:237-44, 1988; Higgins & Sharp, *CABIOS* 5:151-3, 1989; Corpet et al., *Nuc. Acids Res.* 16:10881-90, 1988; Huang et al. *Computer Appls. in the Biosciences* 8, 155-65, 1992; and Pearson et al., *Meth. Mol. Bio.* 24:307-31, 1994. Altschul et al. (*Nature Genetics* 6:119-29, 1994), presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al. *J. Mol. Biol.* 215:403-10, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. It can be accessed at http://www.ncbi.nlm.nih.gov/BLAST/. A description of how to determine sequence identity using this program is available at http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html. As used herein, sequence identity is commonly determined with the BLAST software set to default

parameters. For instance, blastn (version 2.0.6) software may be used to determine sequence identity between two nucleic acid sequences using default parameters. For comparison of two polypeptides, blastp (version 2.0.6) software may be used with default parameters.

An alternative alignment tool is the ALIGN Global Optimal Alignment tool (version 3.0) available from Biology Workbench at http://biology.ncsa.uiuc.edu. This tool may be used with settings set to default parameters to align two known sequences. References for this tool include Meyers and Miller (*CABIOS* 4:11-7, 1989).

Homologs of the disclosed RRV nucleic acids typically possess at least 50% sequence identity counted over the length of one of the nucleic acids (the reference nucleic acid) using the NCBI Blast 2.0.6, gapped blastn set to default parameters. Nucleic acids showing substantial similarity when assessed by this method may show, for example, at least 50%, 60%, 70%, 80%, 90%, 95% or even 98% or greater sequence identity. When less than the entire sequence is being compared for sequence identity, substantially similar nucleotide sequences will typically possess at least 70% sequence identity over short windows of 30-90 nucleic acids, and may possess sequence identities of at least 80%, 90%, 95% or 98% or greater.

Homologs of the disclosed RRV proteins typically possess at least 50% sequence identity counted over full-length alignment with the amino acid sequence of RRV using the NCBI Blast 2.0, gapped blastp set to default parameters. For comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequence will show increasing percentage identities when assessed by this method, such as at least 50%, at least 55%, at least 60%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity. When less than the entire sequence is being compared for sequence identity, homologs will typically possess at least 70% sequence identity over short windows of 10-20 amino acids, and may possess sequence identities of at least 75%, at least 85% or at least 90%, at least 95% or 98% depending on their similarity to the reference sequence. Methods for determining sequence identity over such short windows are described at http://www.ncbi.nlm.nih.gov/BLAST/blast_FAQs.html.

When comparing degrees of sequence identity between similar proteins, the degree of identity will be equal to or less than that the degree of similarity , due to the fact the similarity takes into account conservative amino acid substitutions. So, for instance, the degree of sequence identity between to substantially similar proteins may be at least 43%, 50%, 55%, 65%, 75%, 85%, 95%, 98% or more.

One of ordinary skill in the art will appreciate that these sequence identity ranges are

provided for guidance only; it is entirely possible that strongly significant homologs could be obtained that fall outside of the ranges provided. The present invention provides not only the peptide homologs that are described above, but also nucleic acid molecules that encode such homologs.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions, as described in Example 23.

Specific binding agent: An agent that binds substantially only to a defined target. As used herein, the term "RRV peptide specific binding agent" includes anti-RRV peptide antibodies and other agents that bind substantially only to the RRV peptide. Such "peptide specific binding agents" include anti-IL-6 and anti-MIP antibodies. The antibodies may be monoclonal or polyclonal antibodies that are specific for an RRV peptide, as well as immunologically effective portions ("fragments") thereof. In one embodiment, the antibodies used in the present invention are monoclonal antibodies (or immunologically effective portions thereof) and may also be humanized monoclonal antibodies (or immunologically effective portions thereof). Immunologically effective portions of monoclonal antibodies include Fab, Fab', $F(ab')_2$, Fabc and Fv portions (for a review, see Better and Horowitz, *Methods. Enzymol.* 178:476-96, 1989). Anti-inhibitory peptide antibodies may also be produced using standard procedures described in a number of texts, including Antibodies, A Laboratory Manual by Harlow and Lane, Cold Spring Harbor Laboratory (1988).

Methods of making humanized monoclonal antibodies are well known, and include those described in U.S. Patent Nos. 5,585,089; 5,565,332; 5,225,539; 5,693,761; 5,693,762; 5,585,089; and 5,530,101 and references cited therein. Similarly, methods of making and using immunologically effective portions of monoclonal antibodies, also referred to as antibody fragments, are well known and include those described in Better and Horowitz, 1989, *Meth. Enzymol.* 178:176-496; Better et al., 1990, Better and Horowitz, 1990, Advances in Gene technology: The Molecular Biology of Immune Disease & the Immune Response (ICSU Short Reports); Glockshuber et al., 1990, *Biochemistry* 29:1362-7; and U.S. Patent Nos. 5,648,237; 4,946,778 and 5,455,030, and references cited therein.

The determination that a particular agent binds substantially only to an RRV peptide may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Antibodies, A Laboratory Manual by Harlow and Lane). Western blotting may be used to determine that a given RRV peptide binding agent, such as an anti-IL-6 or MIP peptide monoclonal antibody, binds substantially only to the specific RRV protein.

Subject: This term includes both human and non-human subjects. Similarly, the term "patient" includes both human and veterinary subjects.

Transformed: A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

Transgenic Cell: Transformed cells which contain foreign, non-native DNA.

Variants of Amino Acid and Nucleic Acid Sequences: The production of RRV proteins can be accomplished in a variety of ways (for example see Examples 17, 21 and 25). DNA sequences which encode the protein, or a fragment of the protein, can be engineered such that they allow the protein to be expressed in eukaryotic cells, bacteria, insects, and/or plants. In order to accomplish this expression, the DNA sequence can be altered and operably linked to other regulatory sequences. The final product, which contains the regulatory sequences and the therapeutic protein, is referred to as a vector. This vector can then be introduced into the eukaryotic cells, bacteria, insect, and/or plant. Once inside the cell the vector allows the protein to be produced.

One of ordinary skill in the art will appreciate that the DNA can be altered in numerous ways without affecting the biological activity of the encoded protein. For example, PCR may be used to produce variations in the DNA sequence which encodes RRV proteins, such as IL-6 or MIP. Such variants may be variants that are optimized for codon preference in a host cell that is to be used to express the protein, or other sequence changes that facilitate expression.

Two types of cDNA sequence variant may be produced. In the first type, the variation in the cDNA sequence is not manifested as a change in the amino acid sequence of the encoded polypeptide. These silent variations are simply a reflection of the degeneracy of the genetic code. In the second type, the cDNA sequence variation does result in a change in the amino acid sequence of the encoded protein. In such cases, the variant cDNA sequence produces a variant polypeptide sequence. In order to preserve the functional and immunologic identity of the encoded polypeptide, it is preferred that any such amino acid substitutions are conservative. Conservative substitutions replace one amino acid with another amino acid that has some homology in size, hydrophobicity, etc. Such substitutions generally are conservative when it is desired to finely modulate the characteristics of the protein. For example, conservative substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain.

Examples of amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative substitutions include: Ser for Ala; Lys for Arg; Gln or His for Asn; Glu for Asp; Ser for Cys; Asn for Gln; Asp for Glu; Pro for Gly; Asn or Gln for His; Leu or Val for Ile; Ile or Val for Leu; Arg or Gln for Lys; Leu or Ile for Met; Met, Leu or Tyr

for Phe; Thr for Ser; Ser for Thr; Tyr for Trp; Trp or Phe for Tyr; and Ile or Leu for Val.

The substitutions which in general are expected to produce the greatest changes in protein properties will be non-conservative, for instance changes in which (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histadyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine.

Variations in the DNA sequence that result in amino acid changes, whether conservative or not, should be minimized in order to preserve the functional and immunologic identity of the encoded protein. The immunologic identity of the protein may be assessed by determining whether it is recognized by an antibody to an RRV protein; a variant that is recognized by such an antibody is immunologically conserved. Any DNA sequence variant will preferably introduce no more than 20, and preferably fewer than 10 amino acid substitutions into the encoded polypeptide. Variant amino acid sequences can, for example, be 80%, 90%, 95% or even 98% identical to the native amino acid sequence.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in the host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

Virion: A complete viral particle including envelope, capsid (if any), and nucleic acid elements.

The present invention utilizes standard laboratory practices for the cloning, manipulation and sequencing of nucleic acids, purification and analysis of proteins and other molecular biological and biochemical techniques, unless otherwise stipulated. Such techniques are explained in detail in standard laboratory manuals such as Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor, New York. 1989) and Ausubel et al. (Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences. 1987).

EXAMPLE 1

Isolation of RRV

This example describes how RRV was isolated from a rhesus macaque monkey. Fresh, dispersed bone marrow (BM) cells were isopynic gradient-purified (Ficoll-Paque, Pharmacia) from a 2 yr, 202 day old captive-reared rhesus macaque that was euthanized 503 days after intravenous infection with an SIVmac239 variant. Gradient-purified BM mononuclear cells were seeded into T-25 culture flasks and grown in the presence of Endothelial SFM media (GIBCO) supplemented

with 10% fetal bovine serum, 1% L-glutamine, 1% penicillin-streptomycin-neomycin and 30 μg/mL endothelial cell growth supplement.

Cultures developing cytopathic effects (CPE) were rapidly frozen in liquid N_2 and thawed, and supernatants clarified by centrifugation and filtered through a 0.45 μ membrane. Filtered extracts were then used as inoculum on primary rhesus macaque fibroblast cultures. Fibroblast cultures developing CPE were scraped free into medium, pelleted at 400 x*g*, washed in phosphate-buffered saline and suspended in cold Ito and Karnovsky's fixative (2.5% glutaraldehyde, 0.5% picric acid, 1.6% paraformaldehyde, 0.005% ruthenium red) in 0.1 M sodium cacodylate buffer, pH 7.4 for 2 hours. Fixed cells were washed in cacodylate buffer, post-fixed in 1% OsO_4 and 0.8% $K_3Fe(Cn)_6$ in cacodylate buffer for 1 hour, rinsed in distilled H_2O and pre-stained in 2% aqueous uranyl acetate for 1 hour. Fixed and pre-stained cells were dehydrated in a graded series of acetone imbedded in Epon 812 epoxy resin, polymerized at 60°C and sectioned at 60 nm on an MT 5000 ultramicrotome. Copper grid mounted sections were stained with lead citrate and Uranyl acetate and viewed on a Phillips 300 electron microscope.

By electron microscopy, numerous herpesvirus particles were observed in the cells. This macaque developed LPD characterized as lymphocytic masses in myeloid and nonlymphoid tissues which were confirmed histopathologically as solid pleomorphic lymphoid masses.

EXAMPLE 2

Initial Characterization of RRV

Infectious virus was purified from infected primary rhesus fibroblast cultures exhibiting 100% CPE (see Example 1). Infected cells were harvested and disrupted by freeze-thawing and the cell debris removed by low speed centrifugation. Supernatants were centrifuged in a Beckman JA-14 rotor for 1 hour at 9000 rpm to pellet the virus, which was further purified through a six-step sorbitol gradient ranging from 20 to 70%, spun in a Beckman SW41 rotor for 2 hours at 18,000 rpm. Virus was diluted in balanced buffered salts solution and then pelleted through a 20% sorbitol cushion. Pelleted virus was resuspended in Tris-EDTA buffer (TE; 10 mM Tris-HCl, pH 8.0, and 1 mM EDTA) and lysed in TE with 0.6% SDS and proteinase K (200 μg) at 37°C for 5 hours. Viral DNA was then isolated by $CsCl_2$ gradient centrifugation in a Beckman Ti 75 rotor at 38.4 K rpm for 72 hours, collected and dialyzed against TE.

The viral DNA was analyzed using degenerate primer polymerase chain reaction (PCR) amplification and Southern blot hybridization with a probe specific for the KSHV thymidylate synthase (TS) gene: (5'-CTATACTGCCATTTC-3', SEQ ID NO 180 and 5'-ATGTTTCCGTTTGTA-3', SEQ ID NO 181). The probe itself had the sequence of the KSHV TS (Orf 70 gene). Four genes were identified by these methods. A fragment encoding a portion of the viral DNA polymerase was obtained and DNA sequence analysis revealed that the virus was most likely a gamma herpesvirus, as amino acid sequence identity was highest among this class of

herpesviruses. DNA sequence analysis of the viral DNA fragment found to hybridize to the KSHV TS probe revealed three open reading frames (ORFs) with homology to KSHV (Nicholas et al., 1997, *Nature Med.* 3:287-92; Russo et al.; 1996, *Proc Natl Acad Sci USA* 93:14862-7). One ORF encodes a homologue of macrophage inflammatory protein MIP-1 with amino acid sequence identity with KSHV MIP-II, the second ORF encodes a thymidylate synthase homologue and the third ORF encodes a homologue of interleukin-6 (IL-6) with homology to the rhesus IL-6 and KSHV IL-6. The presence of an IL-6-like cytokine and an MIP-1-like CC-chemokine flanking TS resembles the genomic organization of KSHV, indicating this virus is related to KSHV and is referred to herein as rhesus rhadinovirus (RRV).

To determine if RRV is present in tissue containing the lymphocytic masses, oligonucleotide PCR primers specific for the RRV MIP gene (vMIP-1, 5'-CCTATGGGCTCCATGAGC-3', SEQ ID NO 182; and vMIP-2, 5'-ATCGTCAATCAGGCTGCG-3', SEQ ID NO 183) were designed in an attempt to detect viral DNA in tissue from the macaque. By semi-quantitative PCR analysis, viral DNA sequences were detected in DNA samples from bone marrow at approximately 590 copies per 0.1 μg of tissue DNA. Because rhesus macaques held in captivity are reported to be naturally infected with a herpesvirus similar to KSHV, bone marrow DNA samples were isolated from normal and SIVmac239-infected macaques without LPD and analyzed by PCR. There was no evidence of viral DNA sequences. Additionally, since simian Epstein-Barr virus (EBV) has been found to be present in high copy number in lymphomas from SIV-infected macaques (Baskin et al., 1986, *J. Natl. Cancer Inst.* 77:127-39; Feichtinger et al., 1990,. *Amer. J. Pathol.* 137:1311-5), the tissue samples from the macaque with disease were also analyzed by PCR for rhesus EBV (RhEBV) using oligonucleotide primers for RhEBV latent membrane protein 1. By this analysis, no signal for RhEBV was detected, suggesting that the RRV may be a contributing factor for LPD in this SIV-infected macaque.

EXAMPLE 3

Preparation of RRV DNA for Cloning

Primary rhesus fibroblasts grown in two 850 cm^2 roller bottles were infected with RRV at an MOI of 0.1 and the virus was harvested from the culture supernatant and the infected monolayers 10 to 12 days post-infection. Cellular debris was removed from the culture supernatant by centrifugation at 1,000 x *g* for 10 minutes. Intracellular virus particles were released by sonication followed by centrifugation to pellet debris.

The two clarified supernatants were then combined and the virus was pelleted by centrifugation at 12,500 x *g* for 1 hour at 4°C, and further purified through a six-step sorbitol gradient ranging from 20 to 70%. Gradients were centrifuged in a Beckman SW41 rotor for 2 hours at 18,000 rpm at 4°C. The interface containing the virus was collected and diluted with cold buffered saline solution. The virus was then pelleted by centrifugation in the SW41 for 50

minutes at 18,000 rpm. The virus pellet was resuspended in 9.2 ml of TE (see Example 2) before the addition of 0.6 ml of 10% sodium dodecylsulfate (SDS) and 0.2 ml of proteinase K (10 mg/mL) to release the viral DNA. Viral DNA was isolated by $CsCl_2$ gradient centrifugation in a Beckman Ti75 rotor at 38,400 rpm for 72 hours, collected, and dialyzed against TE.

To ensure that the DNA isolated contained all the necessary sequences required for RRV replication, DNA was transfected, in duplicate, into primary rhesus fibroblasts by the calcium phosphate method without dimethyl sulfoxide shock and observed for cytopathic effects (CPE). Control transfections, lacking viral DNA or calcium phosphate, did not develop CPE.

EXAMPLE 4

Construction of the Cosmid Library

Approximately 100 μg of purified RRV DNA (Example 3) was partially digested with Sau3A I. Aliquots taken at various time points were run on a 0.5% agarose gel and examined for the fraction which gave the desired range of fragments (30 - 42 kb). The selected fraction was dephosphorylated with calf intestinal alkaline phosphatase and 1 μg ligated into the cosmid vector SuperCos 1, prepared essentially as described by the manufacturer (Stratagene, La Jolla, CA). The resulting ligation product was packaged using GigaPack II Gold packaging extract (Stratagene) and grown for the isolation of recombinant cosmids.

Individual recombinant cosmids were grown in 3 ml cultures and the cosmid DNA was isolated by alkaline lysis. Cosmid DNA was digested with EcoR1 and the DNA fragments separated on a 0.8% agarose gel. The separated fragments were transferred to nitrocellulose and probed with various PCR amplification products corresponding to specific KSHV ORFs. Hybridization of the probes to the transferred recombinant cosmids was done under conditions of moderate stringency (2x SSC-0.1%SDS at 55°C) with each of the KSHV-specific probes and at high stringency (0.2x SSC-0.1%SDS at 60°C) with the RRV-specific probes. By this analysis and restriction endonuclease mapping, the recombinant cosmids were aligned and a set of recombinants was identified that represented the entire viral genome when compared to digested viral DNA.

EXAMPLE 5

Cloning and Sequencing

Ten micrograms of each purified recombinant cosmid (Example 4) were digested with EcoRI and the resulting fragments isolated from a 0.8% agarose gel using the QiaQuick gel extraction protocol (Qiagen). Recovered fragments were ligated into pSP73 (Promega). Individual clones were selected by restriction enzyme screening of DNA recovered by alkaline lysis from overnight cultures. Sequencing templates were prepared by alkaline lysis, followed by precipitation with 6.5% polyethylene glycol and 0.8 M NaCl. Templates were resuspended at a concentration of 0.1 μg/μl and end sequences were determined using primers corresponding to the

SP6 and T7 promoters of pSP73. Internal sequences were determined using a combination of subcloning using convenient restriction sites and custom primers. DNA sequencing reactions were performed with the Applied Biosystems (ABI) PRISM Dye Terminator Cycle Sequencing Ready Reaction kits with AmpliTaq DNA polymerase per the manufacturer's instructions. Sequence data was acquired using an ABI 373A Sequencer in the Molecular Biology Core at the Oregon Regional Primate Research Center. The primary EcoRI fragments were sequentially arranged by sequencing across the EcoRI sites in the intact cosmids using custom primers. Except for those regions containing long, high GC repeat units, the entire viral DNA sequence was determined with a redundancy of 3- to 4-fold.

Sequences not accessible to custom primers or restriction subcloning were determined following deletion subcloning using the Exo Size Deletion kit (New England Biolabs). To accommodate this protocol, fragments were subcloned into vectors with restriction sites capable of generating the needed 3' and 5' overhanging ends. Double restriction digests to generate 3' and 5' overhanging ends were performed on 10 μg of recombinant plasmid DNA, which was then subjected to exonuclease III digestion. Aliquots were removed from the exonuclease III digests at empirically-determined time points, frozen on dry ice, then, after all the time points had been collected, incubated for 15 minutes at 65°C to inactivate the enzyme. The DNA was then treated with Mung bean nuclease (MBN) for 30 minutes at 30°C. Prior to addition of 3 μl of MBN to the 12 μl exonuclease III product, the enzyme was diluted 1/25 to reduce nonspecific digestion. Nuclease-treated DNA was recovered using the Wizard prep system (Promega), then incubated for 30 minutes with 2.5 units of T4 DNA polymerase (Life Technologies) and 1 μM dNTPs at 37°C. The final product was ligated overnight with T4 DNA ligase and used to transform competent XL1 blue bacteria. Deletion products were size selected by restriction digests of DNA recovered from 3 ml cultures.

EXAMPLE 6

Assembly of the RRV Sequence, Assignment of ORFs, and Nomenclature

Factura (ABI) and Autoassembler (ABI) were used to assemble the final sequence from individual sequencing runs. Open reading frames in the RRV sequence were determined with the program MacVector (Oxford Molecular Group), using a setting of 100 or more amino acids. Putative ORFs were then translated and compared to a database of KSHV ORFs. RRV ORFs which matched KSHV ORFs were then compared to GenBank using BLASTP to verify the similarity, followed by a Gap analysis (Wisconsin GCG analysis package; Oxford Molecular Group) to determine the levels of similarity and identity between the RRV and KSHV proteins. When a gap in the genome of RRV corresponded to the location of a KSHV ORF with less than 100 amino acids, MacVector was reset to a lower limit. RRV ORFs were assigned the names of HVS ORFs when they showed similarity to KSHV ORFs with the same name.

The nucleotide sequence data from this study have been deposited in the GenBank, EMBL, and DDBJ nucleotide sequence databases under accession number AF083501 (SEQ ID NO 1).

EXAMPLE 7

Primary Structure Of the RRV Genome

The genomic nucleotide sequence of the RRV genome (as shown in SEQ ID NO 1) was determined using twenty-nine EcoRI fragments (as shown in FIG. 2) from seven overlapping isolates of a partial Sau3A I cosmid library. Cosmids were selected by hybridization with PCR products from KSHV ORFs. EcoRI fragments from each cosmid were subcloned into pSP73 (Promega) and sequenced. The EcoRI fragments were arranged in the proper order by sequencing across the EcoRI junctions in the parent cosmids using custom primers. Greater than 98% of the viral genome was determined on both strands. The average sequencing redundancy was between 3 and 4, but three regions were sequenced on only one strand. One of these regions is a 106 bp segment of ORF 61 (SEQ ID Nos 136 and 137) that was blocked on one side by an apparent hairpin. This segment was sequenced multiple times in one direction using templates derived from independent overlapping cosmids. The other two regions are 1 kb, high G + C, repetitive sequences. These segments, which are discussed in more detail below, were sequenced completely on one strand using a combination of custom primers and exonuclease III deletions.

Terminal repeats were identified on both the left and right ends of the genome and the sequence between them was designated as the LUR of the genome. The first base to the right of the left terminal repeat was designated base one. The LUR is 133,719 bp long (SEQ ID NO 1). The G + C content of RRV is 52.2%, which is comparable to the 53.5% G + C content of KSHV, but considerably higher than the 34.5% G + C content of the HVS genome. The CpG ratio is 1.11, which is substantially higher than the ratio found for other gamma-herpesviruses.

ORFs were identified by MacVector and compared to a database containing the full complement of known KSHV ORFs. Matches between RRV and KSHV proteins were verified by a BLASTP search of GenBank with the RRV proteins and then by Gap analysis. The initial screening for ORFs used a minimum size limit of 100 amino acids. This limit was reduced when smaller KSHV ORFs existed in locations corresponding to unassigned regions of RRV. Using this approach, 82 ORFs were identified (even-numbered SEQ ID Nos 2-164), with 67 of these corresponding to ORFs found in both KSHV and HVS. In accordance with the standard nomenclature for rhadinoviruses, these ORFs were labeled according to the HVS designation. The 15 ORFs not found in HVS were assigned labels beginning with R (for rhesus), indicating their presence in RRV, but not HVS. Some of these genes have counterparts in KSHV.

A map of the genome of RRV is presented in FIG. 3, with all identified ORFs and their orientations. The BamHI, EcoRI, and HindIII restriction sites in relation to the genome are shown

in FIG. 2. The BamHI and Hind III maps were generated from the final compiled sequence. The EcoRI map was also generated from the final compiled sequence, but it was further characterized by sequencing across the EcoRI junctions in the parent cosmids. Fragment sizes for each restriction map are presented in FIG. 4.

EXAMPLE 8

Genomic Organization of RRV

The overall genomic organization of RRV matches the general structure of gamma-herpesviruses, with blocks of shared ORFs interrupted at specific locations (referred to as divergent loci) where the viral genomes code for acquired cellular genes. The primate rhadinoviruses form a subset of the gamma-herpesviruses and their genomes are correspondingly more similar to each other than to other members of the family.

The genomic sequence of RRV is presented in SEQ ID NO 1. FIG. 3 shows a schematic representation of the ORFs of RRV with a corresponding restriction map. FIG. 4 shows the location, size and description of the RRV ORFs.

EXAMPLE 9

Comparison of RRV and KSHV ORFs

A comparison of corresponding repeats in RRV and KSHV is shown in FIG. 5. In addition, FIG. 5 presents data for RRV ORFs along with the results of the Gap analysis of ORFs shared by RRV, KSHV, and HVS. All HVS-like ORFs found in KSHV are found in RRV. A comparison table of interferon regulatory elements encoded by the RRV and KSHV genomes is shown in FIG. 6.

EXAMPLE 10

Comparison of RRV and HVS ORFs

FIG. 7 shows the results of the Gap analysis of ORFs shared by RRV, KSHV, and HVS. In general, RRV and HVS ORFs are highly similar when the corresponding RRV and KSHV ORFs are highly similar, although the Gap values are generally lower.

EXAMPLE 11

ORFs Unique to RRV and KSHV

RRV includes 14 genes which are not found in HVS (R1 SEQ ID NOS 2 and 3; R2 SEQ ID NOS 20 and 21; R3 SEQ ID NOS 24 and 25; R4 SEQ ID NOS 98 and 99; R5 SEQ ID NOS 100 and 101; R6 SEQ ID NOS 114 and 115; R7 SEQ ID NOS 116 and 117; R8 SEQ ID NOS 118 and 119; R9 SEQ ID NOS 120 and 121; R10 SEQ ID NOS 122 and 123; R11 SEQ ID NOS 124 and 125; R12 SEQ ID NOS 126 and 127; R13 SEQ ID NOS 128 and 129; and R15 SEQ ID NOS 160 and 161). These are designated in FIG. 3 as “R” ORFs. Of these fifteen genes, 11 have

counterparts in the genome of KSHV. R2 (SEQ ID NOs 20 and 21) and R3 (SEQ ID NOs 24 and 25) are cytokine genes. R2 has functional homology to K2, the vIL-6 gene of KSHV. Gap analysis of the vIL-6 genes from KSHV and RRV shows no notable similarity, but both possess four conserved cysteines found in cellular IL-6. In addition, RRV vIL-6 has IL-6-like activity in cell culture. R3 has a small, but clear, similarity to KSHV K4, a vMIP1β gene. It is the only vMIP gene in RRV, as compared to the three vMIP genes found in KSHV.

RRV R6 through R13 are vIRFs as are KSHV K9 through K11 (FIG. 6). K9, the most studied of the KSHV vIRFs, does not have a DNA binding domain, but has been demonstrated to inhibit the endogenous cellular interferon response pathways. Five of the RRV vIRFs (R6, R7, R8, R10, and R11) are similar to K9, though only R10 has a similarity greater than 30%. The remaining similarities fall between 26% and 30%. There is no measurable similarity between any RRV vIRF and any KSHV vIRF other than K9. There is, however, a pattern of higher similarity between members of the RRV vIRF family. R6, R7, R8, and R9 are most similar to R10, R11, R12, and R13, respectively, with the similarities falling between 50% and 62%. The pattern of similarity suggests a single, possibly recent, gene duplication event for RRV which increased the number of vIRFs in the genome from four to eight.

The final RRV gene with a unique KSHV counterpart is R15, which has some similarity to K14, a viral NCAM Ox-2 homologue. The similarity between R15 and K14 (35.2%) is relatively low compared to most other shared proteins.

A number of genes in RRV appear to be truly unique. R1 colocalizes with, but has no similarity to, K1, a KSHV gene that has been demonstrated to have *in vivo* transforming ability. K1 and R1 both colocalize with ORF1, or STP (saimiri transforming protein), although both K1 and R1 are in opposite orientations compared to STP. A BLASTP search of GenBank using R1 reveals a limited amino-terminal similarity to a series of Fc receptors, including a potential transmembrane domain. These data suggest that R1, like K1 and STP, may have transforming potential via transmembrane signaling.

R4 and R5 are located between ORF 50 and 52, the same location as K8 and K8.1 in KSHV; however, there is no similarity between either R4 or R5 and the KSHV proteins. A BLASTP search of GenBank failed to show any significant alignments with either R4 or R5, so their functions are unknown.

RRV has no confirmed ORFs in the region corresponding to K12, the ubiquitously expressed kaposin gene. A large ORF exists to the right of ORF 71, but it has no apparent control regions (TATA box or polyadenylation signal), so it has not been designated as a true ORF, pending identification of transcripts from this region. No ORFs corresponding to KSHV K15 have been identified.

EXAMPLE 12

Co-localization of Repeat Units in RRV and KSHV

The RRV genome contains three highly repetitive regions, which correspond to three of the repetitive regions of KSHV: *frnk*, *zppa*, and *mdsk* (FIG. 5). KSHV *frnk* and *zppa*, and the corresponding RRV repetitive regions, *rDL-B* and *rDL-E*, respectively, are tandem repeats.

The first element of the RRV *syko* repeat is much lower in G + C content than the corresponding KSHV element, although the sizes are comparable (FIG. 5). The second element is over 700 bp longer than the corresponding KSHV element. The first element of the RRV *vrtgo* repeat is 30% longer than the corresponding KSHV element, and the second RRV element is over four times as long as the second KSHV element. There is no sequence similarity between the various elements of the two viruses nor is there any similarity between any two repeat sequences in RRV.

Not all repeat elements found in KSHV have corresponding repeats in RRV. This includes the KSHV *vnct* and *waka/jwka* repeats. This also includes the *moi* repeat, which is located in the center of the KSHV ORF 73 and is responsible for the divergent lengths of RRV and KSHV ORF 73. *Moi* is described in the annotations to the KSHV GenBank entry as having 15 different 11-16 bp repeats. The result of this repeat element is the presence in ORF 73 of a highly acidic central domain, with a large number of glutamate residues coded by a repeating GAG codon. KSHV ORF 73 is a potential leucine zipper protein, with a number of leucine zipper sites in the repeat region. RRV lacks the *moi* repeat and its concomitant acidic domain. It also lacks any evidence for a leucine zipper, indicating that the biology of ORF 73 in RRV may be substantially different than the biology of ORF 73 in KSHV.

EXAMPLE 13

Production of Simian Kaposi's Sarcoma (KS) and Lymphoproliferative Disorders Model

This example describes how the RRV cloned above can be used to produce a non-human primate model for Kaposi's sarcoma and lymphoproliferative disorders. Four rhesus macaques (identification numbers 18483, 18503, 18540 and 18570) that were approximately 1.5 years old, and PCR- and seronegative for RRV were selected. To perform the antibody analysis, infected cells were solubilized with 0.5% Nonidet P-40 and 1% sodium deoxycholate in phosphate buffered saline, and clarified in a Beckman SW28 rotor at 23,500 rpm for 1 hour at 4°C. The clarified supernatant was used as antigen for coating enzyme-linked immunosorbent assay (ELISA) plates (500 ng/well). ELISAs were then performed essentially as described by Kodama et al. (*AIDS Res Hum Retroviruses* 5:337-43, 1989).

All of the animals were then inoculated intravenously with cell-free supernatants containing the equivalent of 5 ng of p27 prepared from COS-1 cells transfected with an

SIVmac239 molecular clone (Endres et al., 1995, *SW. J Med. Primatol.* 24:141-4). The PBMCs from all macaques were prescreened for *in vitro* susceptibility to virus infection as described by Naidu et al. (*J. Virol.* 62:4691-6, 1988). All inoculations and animal manipulations were performed according to institutional guidelines at the Oregon Regional Primate Research Center (Beaverton, OR). Every 3-4 days for 4 weeks, then at 2-week intervals, macaques were sedated with ketamine hydrochloride (10 mg per kilogram of body weight) and examined for fever, weight loss, cutaneous signs, lymphadenopathy, and hepatomegaly or splenomegaly. At these times, venipuncture was performed and blood specimens collected. Plasma was monitored for virus during the first 4 weeks with the SIV p27 enzyme-linked immunosorbent assay (ELISA) (Coulter Corp. Hialeah, FL.). T cell subsets and B cells were measured by flow cytometry with the OKT4 (CD4, Ortho), B9.11 (CD8, Coulter), and B-Ly-1 (CD20, Coulter) monoclonal antibodies.

At 8 weeks post-SIV infection, rhesus macaques 18483 and 18570 were inoculated intravenously with 5 x 10^6 plaque forming units of gradient purified RRV that was grown and titered by plaque assay on primary rhesus fibroblasts. The two remaining macaques (18503 and 18540) were kept as SIV-infected controls. Every 3-4 days for 2 weeks, once a week for 4 weeks, then at 2 week intervals, the macaques were examined and blood samples collected and analyzed. Virus isolations were performed by cocultivation of 2 x 10^5 PBMCs from each of the macaques with primary rhesus fibroblasts in duplicate. Cell cultures were monitored every 2-3 days for 3-4 weeks for cytopathic effects characteristic of RRV. PBLs were also analyzed by PCR for the presence of viral DNA. PCR analysis for RRV was performed with the following oligonucleotide primers: vMIP-1, 5' CCTATGGGCTCCATGAGC 3' (SEQ ID NO 166); and vMIP-2, 5' ATCGTCAATCAGGCTGCG 3' (SEQ ID NO 167). The conditions for PCR were 94°C for 2 minutes (1 cycle); 94°C for 0.5 minutes, 50°C for 0.5 min, 72°C for 0.5 minutes (30 cycles); 72°C extension for 5 minutes (1 cycle). Each PCR reaction used 0.1 Fg of total DNA, 50 pmole of each primer, 1 U of Vent polymerase, 40 μM each of deoxynucleotide triphosphate, 10 mM KCl, 10 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$ and 0.1% Triton X-100 in a final volume of 50 μL. The PCR reactions were run out on a 1% agarose gel, transferred to nitrocellulose, and probed with a ^{32}P-ATP-labeled oligonucleotide primer specific for vMIP-3 (5' ATATTAAACACTCGCCGC- 3' SEQ ID NO 168). Hybridizations were performed overnight at room temperature in 6X SSC, 0.1% SDS and 10 μg/mL *E. coli* tRNA. Southern blots were then washed with 2X SSC and 0.1% SDS twice at room temperature followed by two washes for 1 hour in 2X SSC and 0.1% SDS at 47°C. Bound probe was visualized by exposing NEN duPont reflection film to the washed membrane at 80°C with an NEN duPont Reflection screen.

Infectious RRV was recovered from the peripheral blood mononuclear cells (PBMCs) of both RRV macaques injected with RRV as early as 4 weeks after inoculation for one macaque (18570) and 8 weeks for the other macaque (18483), but not from the control macaques. The peripheral blood leukocytes (PBL) from both macaques were also shown to harbor viral DNA as

determined by PCR and Southern blot analysis for the viral MIP gene, as early as 4 weeks after inoculation for one macaque (18483) and as late as 14 weeks for the second macaque (18570). Additionally, antibody responses to RRV were observed as measured by ELISA in the RRV-infected macaques beginning 4 weeks post-infection, but not in the control macaques.

Flow cytometry analysis (FACS) of PBLs at the indicated weeks post-infection (FIGS 8A-8D) showed there was limited CD4+ lymphocyte depletion after SIV infection in both groups of macaques followed by a rebound and sustained CD4+ lymphocyte counts. However, examination of CD20+ B lymphocytes revealed significant differences between the two groups. The two control macaques exhibited a dramatic and sustained decline in CD20+ B lymphocytes (FIGs. 8C and 8D), whereas both co-infected macaques exhibited a transient increase in B lymphocytes beginning 6 weeks after RRV infection (FIGs. 8A and 8B). The increase in CD20+ B lymphocytes correlated with the isolation and/or detection of RRV in both macaques; however, viral load did not appear to correlate with the increase in CD20+ B lymphocytes when all samples from each macaque were analyzed simultaneously. It has been reported that CD23, a B cell activation marker, is induced by RhEBV infection of macaques (Moghaddam et al. 1997, *Science* 276:2020-33). FACS analysis of PBMCs from RRV-infected macaques revealed no detectable CD23+ cells. This would suggest that the mechanism responsible for increased numbers of CD20+ B lymphocytes following RRV infection differ from the activation of B lymphocytes by RhEBV.

Routine physical examinations were performed on all four macaques, and early symptoms of SIV infection were observed in all four macaques by 2 weeks, including fever, rash and malaise. However, 11 weeks after inoculation with RRV, macaques 18483 and 18570 developed marked lymphadenopathy and splenomegaly, estimated to be enlarged 10 to 20 times the size of a normal spleen. In contrast, there was only slight lymph node enlargement in the control macaques not infected with RRV and no detectable enlargement of the spleen. Lymph node biopsies of the RRV-infected macaques revealed almost identical histology, characterized by a predominately follicular lesion with giant germinal centers and paracortical hyperplasia with increased vascularity, resembling angiofollicular lymph node hyperplasia associated with KSHV in Castleman's disease (Lachant et al. 1985, *Am. J. Clin. Pathol.* 83:27-33). In contrast, the lymph nodes of the control macaques exhibited atrophied lymphoid follicles and paracortical depletion characteristic of SIV-induced lymphoid atrophy (Chalifoux et al., 1987, *Am. J. Pathol.* 128:104-10; Ringler et al., 1989, *Am. J. Pathol.* 134:373-83; Wyand et al, 1989, *Am. J. Pathol.* 134:385-93). By FACS analysis, the majority of the lymph node mononuclear cells were CD20+ B lymphocytes in RRV-infected macaques, whereas CD4+ and CD8+ T lymphocytes predominated in the control macaques.

The presence of viral DNA was determined by PCR analysis on DNA derived from PBLs. Detection of antibodies to RRV was determined by enzyme-linked immunosorbent assay (ELISA)

on plates coated with extracts derived from RRV-infected cells. By PCR analysis, RRV sequences were more prevalent in the lymph nodes than in the peripheral blood of RRV-infected macaques, whereas control macaques were negative for RRV sequences (FIGs. 9A and 9B).

Additional disease manifestations were also observed in the RRV-infected macaques that parallel clinical features and B cell abnormalities observed in AIDS patients. Hypergammaglobulinemia was observed in the RRV-infected macaque that the virus was derived from, as well as in the macaques experimentally infected with RRV, whereas the two control macaques had gammaglobulin levels similar to those before SIV infection. In addition, one of two RRV-infected macaques (18570) developed severe autoimmune hemolytic anemia 30 weeks after RRV infection, a condition frequently observed in MCD patients (Parravicini et al., 1997, *Am. J. Pathol.* 151:1517-22).

The second of the two RRV-infected macaques developed other unique clinical manifestations that paralleled those of AIDS patients with KS. At 60 weeks post-RRV infection it developed a distended abdomen that was clinically evident upon physical examination. Palpation revealed a pronounced fluid accumulation in the peritoneal cavity. This animal was euthanized due to persistent fluid accumulation and hyperbilirubinemia. Necropsy analysis on this animal revealed an abundance of ascites fluid, which was comprised predominately of CD20 B cells, as identified by FACS analysis. In addition, this animal exhibited a mesenchymal proliferative lesion throughout the viscera, that was identified by histopathological examination to be retroperitoneal fibromatosis (RF). RF is an abnormal highly vascularized mesenchymal proliferative lesion that exhibits histological features resembling Kaposi's Sarcoma. Analysis of DNA isolated from the ascites and RF lesion by PCR with RRV MIP primers (given in Example 2) revealed a high viral load, implying RRV infection was responsible for these abnormal proliferations.

EXAMPLE 14

Other Methods to Prepare RRV Nucleic Acid Sequences

Obtaining the RRV Viral Genome

The RRV genome of the invention (SEQ ID NO 1) can be procured by *de novo* isolation from a viral culture. A biological sample of the virus (accession number VR-2601) may be obtained from the ATCC in Manassas, VA. This virus can be grown *in vitro* using primary rhesus fibroblasts (see Example 1). The virus is harvested from the culture supernatant and the infected host cells. Cellular debris is removed by centrifugation and intracellular virus particles may be released by sonication followed by centrifugation to pellet debris. The virus is then pelleted by centrifugation and further purified through a six-step sorbitol gradient. The interface containing the virus is collected and the virus then pelleted by centrifugation, and the viral DNA released by SDS disruption. Viral DNA may be isolated by $CsCl_2$ gradient centrifugation.

Obtaining Selected Polynucleotides from the Viral genome

The isolated viral genome can be used as a source of polynucleotides as identified by the sequence as disclosed herein (SEQ ID NO 1). The polymerase chain reaction (PCR) may be used to amplify any polynucleotide selected from the known viral sequence using the viral genome as a source of template DNA. The template DNA may also be provided in the form of one or more cosmids that contain fragments of the viral genome. Alternately, cDNA, produced by reverse transcription of RNA extracted from RRV infected host cells, may be used as a template in a reverse-transcription PCR (RT-PCR) reaction. Methods and conditions for PCR and RT-PCR amplification are described in Innis et al. (PCR Protocols, A Guide to Methods and Applications, 1990, Innis et al. (eds.), 21-27, Academic Press, Inc., San Diego, California).

The selection of PCR primers may be made according to the portions of the genome to be amplified. Primers may be chosen to amplify small fragments of the genome, ORFs or fragments including many contiguous genes from the genome. Variations in amplification conditions may be required to accommodate primers of differing lengths, and such considerations are well known in the art and are discussed in Innis at al. (PCR Protocols, A Guide to Methods and Applications, 1990, Innis et al. (eds.), 21-27, Academic Press, Inc., San Diego, California), Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor, New York. 1989) and Ausubel et al. (Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences. 1987). For example, the ORF corresponding to the MIP gene may be amplified from an RRV genomic (or appropriate cosmid) template using the following pair of primers: 5' ATGAGGGGCCTTTTCGTGTGC 3' (SEQ ID NO 169) and 5' CTGAATCCCGCTGCCAAGGCC 3' (SEQ ID NO 170).

Likewise, the ORF corresponding to the IL-6 gene may be amplified from an RRV genomic (or appropriate cosmid) template using the following pair of primers: 5' ATGTTCCCTGTCTGGTTCGTC 3' (SEQ ID NO 171) and 5' TTACATCATAGCTATTGCGCG 3' (SEQ ID NO 172).

Such primers are illustrative only and it will be readily appreciated by one of ordinary skill in the art that many different primers may be selected from the sequence disclosed and used in PCR amplification reactions to amplify DNA sequences of interest from the RRV genome.

Polynucleotides that may be obtained by the above methods include, for example: the entire polynucleotide genome of RRV as shown in SEQ ID NO 1; ORFs of this genome; oligonucleotides comprising at least 15, 20, 30, 40, 50, 70, 100 and 150 consecutive nucleotides of the genome sequence as shown in SEQ ID NO 1; nucleic acid sequences defined by nucleotides 1 to 11031 of SEQ ID NO 1 and nucleotides 21625 to 133719 of SEQ ID NO 1; and ORFs selected from these nucleic acid sequences. It is readily apparent that fragments of any length may be made using the above methods and information.

EXAMPLE 15

Therapeutic and Diagnostic Uses of the RRV IL-6 Protein

As disclosed herein, the genome of RRV possesses an IL-6 gene (FIGS. 1 and 10 and SEQ ID NO 20) similar to that found in KSHV. The IL-6 and MIP proteins of KSHV are thought to be important in disease pathology, such as in Kaposi's sarcoma. The primary structure of the RRV IL-6 protein is shown in FIG. 10 (SEQ ID NO 21). Given this sequence information, one can readily make derivative proteins of RRV IL-6. Such derivative proteins include proteins that differ from the primary amino acid sequence as shown in FIG. 10 (SEQ ID NO 21) by one or more conservative amino acid substitutions. Examples of such conservative substitutions are given in the DEFINITIONS section of the specification. Derivative proteins also include proteins consisting of an amino acid sequence that has a defined degree of amino acid similarity with the RRV IL-6 protein. For instance, such derivative proteins will typically have at least 50% sequence similarity (and may have at least 60%, 70%, 80%, 90%, 95%, 98% or even 99% sequence similarity) with the RRV IL-6 protein. Such derivative proteins will not only share sequence similarity with KSHV IL-6 but will also possess IL-6 biological activity.

IL-6 is a cytokine known to have pleiotropic immunological effects including anti-inflammatory and immunosuppressive effects, and may be used in several therapeutic and diagnostic applications. RRV IL-6 of the invention may be likewise be used. For instance, IL-6 may be used to induce stimulation of hematopoietic stem cells, and to encourage proliferation, differentiation and terminal maturation of erythroid cells from hematopoietic cells. Thus, for instance, RRV IL-6 may be used *in vivo* or *ex vivo* to treat diseases that involve leukopenia and thrombocytopenia. Such uses include stimulation of hematopoietic cells of radiotherapy patients or people exposed to radiation accidentally. IL-6 may be used in such applications in conjunction with GM-CSF (granulocyte-macrophage stimulating factor) (see U.S. Patent Nos. 5,610,056 and 5,620,685, herein incorporated by reference). IL-6 can also be used to stimulate growth of megakaryocytes and platelets, and for the inhibition of tumor growth (see U.S. Patent No. 5,620,685, herein incorporated by reference). IL-6 can also be used for the treatment of leukemias, such as chronic myeloid leukemia (CML) and acute myeloid leukemia, by inducing terminal differentiation of cells with IL-6 (see WO 90/01943, herein incorporated by reference). RRV IL-6 may be used for all such applications.

Therapeutic applications may involve the administration of RRV IL-6 in a number of ways. RRV IL-6 may be administered *in vivo*, e.g., by injection systemically or locally, for instance, into a subject. Many other forms of *in vivo* administration are possible including intravenous, subcutaneous, across a mucous membrane (anally, vaginally or sublingually), transdermal or by direct injection. Additionally, it may be administered *ex vivo*, by the removal of cells from a subject, the treatment of these cells *in vitro* with RRV IL-6, and the replacement of

these cells into the subject. Another recently developed method of delivery of a protein drug is by introducing the gene coding for the drug into a subject, for instance within the genome of a virus, such as an adenovirus or a retrovirus, whereby the protein is expressed in the subject. Other modes of administration are provided in Example 25.

Such examples are provided for illustrative purposes only and it will be seen that RRV IL-6 may be used in a variety of topical and systemic immunological treatments where it would be desirable to stimulate cell proliferation or to induce anti-inflammatory or immunosuppressive effects. Additionally, IL-6 of the invention may be used for research and diagnostic purposes as discussed generally herein. For instance, IL-6 may be used to produce antibodies for diagnostic purposes to diagnose diseases characterized by increased or decreased production of IL-6, and the nucleic acid sequence encoding IL-6 may be used to produce probes and primers for diagnostic and research purposes or for gene therapy applications. The IL-6 could also be used as a targeting molecule for identifying cells with receptors for IL-6, and for directing therapeutic agents to these cells, for example by linking detector or therapeutic molecules to IL-6.

EXAMPLE 16

Therapeutic and Diagnostic Uses of the RRV MIP Protein

The genome of RRV as disclosed herein possesses an MIP gene (FIGS. 1 and 11 and SEQ ID NO 24) similar to that found in KSHV. The primary structure of the RRV MIP protein is shown in FIG. 11 (SEQ ID NO 25). Given this sequence information, one can readily make derivative proteins of RRV MIP. Such derivative proteins include proteins that differ from the primary amino acid sequence as shown in FIG. 11 (SEQ ID NO 25) by one or more conservative amino acid substitutions. Derivative proteins also include proteins consisting of an amino acid sequence that has a defined degree of amino acid similarity with the RRV MIP protein. Typically such derivative proteins will have at least 50% sequence similarity with the RRV MIP protein, and may have at least 60%, 70%, 80%, 90%, 95%, 98%, or even 99% sequence similarity. Such derivative proteins will not only share sequence similarity with KSHV MIP but will also possess MIP biological activity. MIP biological activity can be detected and quantified using bioassays as described in Kedal et al. (*Science* 277:1656-9, 1997) and Boshoff et al. (*Science* 278:290-4, 1997) that measure MIP concentrations using HIV inhibition and calcium mobilization, respectively.

MIP is a cytokine that activates neutrophils to undergo an oxidative burst and is also intrinsically pyrogenic. The MIP genes and proteins of the invention may be used in several therapeutic and diagnostic ways. The RRV MIP protein may be used for the same applications as other MIP proteins. Treatment of wounds to promote healing by application of MIP to the wound site is discussed in U.S. Patent No. 5,145,676. U.S. Patent No. 5,474,983 (herein incorporated by reference) discusses various methods of treatment of inflammatory diseases including asthma,

allergies and dermatitis. U.S. Patent No. 5,656,724 (herein incorporated by reference) discloses the use of MIP to suppress proliferation of dividing myeloid cells e.g., for the treatment of neutropenia. Use of MIP to inhibit HIV is discussed by Kedal et al. (*Science* 277:1656-9, 1997). RRV MIP may be used for all such applications.

As illustrated for IL-6 above, MIP may be administered in various ways to provide a therapeutic effect including *in vivo*, *ex vivo* and by gene therapy.

Such examples are provided for illustrative purposes only and it will be seen that MIP may be used in a variety of topical, systemic, *in vivo* and *ex vivo* immunological treatments where it would be desirable to activate neutrophils or to induce fever. Additionally, MIP of the invention may be used for diagnostic purposes as discussed generally herein. For instance, MIP may be used to produce antibodies for diagnostic purposes to diagnose diseases characterized by increased or decreased production of MIP, and the nucleic acid sequence encoding MIP may be used to produce probes and primers for diagnostic and research purposes, or for gene therapy applications.

The MIP could also be used as a targeting molecule for identifying cells with receptors for MIP, and for directing therapeutic agents to these cells, for example by linking detector or therapeutic molecules to MIP.

Although Examples 15 and 16 provide examples of therapeutic uses of the RRV IL-6 and MIP proteins, any of the other proteins encoded by the RRV can also be administered therapeutically, or diagnostically. For example, RRV proteins that induce pathological or physiological conditions in a recipient can be administered to stimulate that condition for study, or to provide an animal or human model of the condition. That model can then be used to study the condition, or treatments for it.

EXAMPLE 17

Expression of RRV cDNA Sequences

With the provision of the RRV genomic (SEQ ID NO 1) and cDNAs (even-numbered SEQ ID Nos 2-164), the expression and purification of any of the RRV proteins (odd-numbered SEQ ID Nos 3-165), from any species, by standard laboratory techniques is now enabled. Fragments amplified as described herein can be cloned into standard cloning vectors and expressed in commonly used expression systems consisting of a cloning vector and a cell system in which the vector is replicated and expressed. Purified proteins may be used for functional analyses, antibody production, diagnosis, and patient therapy. Furthermore, the DNA sequences of the RRV cDNAs can be manipulated in studies to understand the expression of RRV genes and the function of their products. Mutant forms of RRV may be isolated based upon information contained herein, and may be studied in order to detect alteration in expression patterns in terms of relative quantities, and functional properties of the encoded mutant RRV protein. Partial or full-length cDNA sequences, which encode for the protein, may be ligated into bacterial expression vectors. Methods for expressing large amounts of protein from a cloned gene introduced into *E. coli* may

be utilized for the purification, localization and functional analysis of proteins. For example, fusion proteins consisting of amino terminal peptides encoded by a portion of the *E. coli* lacZ or trpE gene linked to RRV protein may be used to prepare polyclonal and monoclonal antibodies against this protein. Thereafter, these antibodies may be used to purify proteins by immunoaffinity chromatography, in diagnostic assays to quantitate the levels of protein and to localize proteins in tissues and individual cells by immunofluorescence and microscopy.

Intact native protein may also be produced in *E. coli* in large amounts for functional studies. Standard prokaryotic cloning vectors may also be used, for example *pBR*322, *pUC*18 or *pUC*19 as described in Sambrook et al. (Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor, New York. 1989). Nucleic acids of RRV may be cloned into such vectors which may then be transformed into bacteria such as *E. coli* which may then be cultured so as to express the protein of interest. Other prokaryotic expression systems include, for instance, the arabinose-induced pBAD expression system that allows tightly controlled regulation of expression, the IPTG-induced pRSET system that facilitates rapid purification of recombinant proteins and the IPTG-induced pSE402 system that has been constructed for optimal translation of eukaryotic genes. These three systems are available commercially from Invitrogen and, when used according to the manufacturer's instructions, allow routine expression and purification of proteins.

Methods and plasmid vectors for producing fusion proteins and intact native proteins in bacteria are described in Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989, Chapter 17). Such fusion proteins may be made in large amounts, are easy to purify, and can be used to elicit antibody response. Native proteins can be produced in bacteria by placing a strong, regulated promoter and an efficient ribosome binding site upstream of the cloned gene. If low levels of protein are produced, additional steps may be taken to increase protein production; if high levels of protein are produced, purification is relatively easy. Suitable methods are presented in Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989) and are well known in the art. Often, proteins expressed at high levels are found in insoluble inclusion bodies. Methods for extracting proteins from these aggregates are described by Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989, Chapter 17).

Vector systems suitable for the expression of lacZ fusion genes include the pUR series of vectors (Ruther and Muller-Hill, 1983, *EMBO J.* 2:1791), pEX1-3 (Stanley and Luzio, 1984, *EMBO J.* 3:1429) and pMR100 (Gray et al., 1982, *Proc. Natl. Acad. Sci. USA* 79:6598). Vectors suitable for the production of intact native proteins include pKC30 (Shimatake and Rosenberg, 1981, *Nature* 292:128), pKK177-3 (Amann and Brosius, 1985, *Gene* 40:183) and pET-3 (Studiar and Moffatt, 1986, *J. Mol. Biol.* 189:113). The RRV fusion protein may be isolated from protein gels, lyophilized, ground into a powder and used as an antigen. The DNA sequence can also be transferred to other cloning vehicles, such as other plasmids, bacteriophages, cosmids, animal

viruses and yeast artificial chromosomes (YACs) (Burke et al., 1987, *Science* 236:806-12). These vectors may then be introduced into a variety of hosts including somatic cells, and simple or complex organisms, such as bacteria, fungi (Timberlake and Marshall, 1989, *Science* 244:1313-7), invertebrates, plants (Gasser and Fraley, 1989, Science 244:1293), and mammals (Pursel et al., 1989, *Science* 244:1281-8), which cell or organisms are rendered transgenic by the introduction of one or more heterologous RRV DNAs.

Various yeast strains and yeast-derived vectors are commonly used for expressing and purifying proteins, for example, *Pichia pastoris* expression systems are available from Invitrogen (Carlsbad, CA). Such systems include suitable *Pichia pastoris* strains, vectors, reagents, transformants, sequencing primers and media.

Non-yeast eukaryotic vectors can also be used for expression of the RRV proteins. Examples of such systems are the well known Baculovirus system, the Ecdysone-inducible mammalian expression system that uses regulatory elements from *Drosophila melanogaster* to allow control of gene expression, and the Sindbis viral expression system that allows high level expression in a variety of mammalian cell lines. These expression systems are available from Invitrogen.

For expression in mammalian cells, the cDNA sequence may be ligated to heterologous promoters, such as the simian virus SV40, promoter in the pSV2 vector (Mulligan and Berg, 1981, *Proc. Natl. Acad. Sci. USA* 78:2072-6), and introduced into cells, such as monkey COS-1 cells (Gluzman, 1981, *Cell* 23:175-82), to achieve transient or long-term expression. The stable integration of the chimeric gene construct may be maintained in mammalian cells by biochemical selection, such as neomycin (Southern and Berg, 1982, *J. Mol. Appl. Genet.* 1:327-41) and mycophoenolic acid (Mulligan and Berg, 1981, *Proc. Natl. Acad. Sci. USA* 78:2072-6).

DNA sequences can be manipulated with standard procedures such as restriction enzyme digestion, fill-in with DNA polymerase, deletion by exonuclease, extension by terminal deoxynucleotide transferase, ligation of synthetic or cloned DNA sequences, site-directed sequence-alteration via single-stranded bacteriophage intermediate or with the use of specific oligonucleotides in combination with PCR.

The cDNA sequence (or portions derived from it) or a mini gene (a cDNA with an intron and its own promoter) may be introduced into eukaryotic expression vectors by conventional techniques. These vectors are designed to permit the transcription of the cDNA eukaryotic cells by providing regulatory sequences that initiate and enhance the transcription of the cDNA and ensure its proper splicing and polyadenylation. Vectors containing the promoter and enhancer regions of the SV40 or long terminal repeat (LTR) of the Rous Sarcoma virus and polyadenylation and splicing signal from SV40 are readily available (Mulligan and Berg, 1981, *Proc. Natl. Acad. Sci. USA* 78:2072-6; Gorman et al., 1982, *Proc. Natl. Acad. Sci USA* 78:6777-81). The level of expression of the cDNA can be manipulated with this type of vector, either by using promoters

that have different activities (for example, the baculovirus pAC373 can express cDNAs at high levels in *S. frugiperda* cells (Summers and Smith, 1985, Genetically Altered Viruses and the Environment, Fields et al. (Eds.) 22:319-328, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.) or by using vectors that contain promoters amenable to modulation, for example, the glucocorticoid-responsive promoter from the mouse mammary tumor virus (Lee et al., 1982, Nature 294:228). The expression of the cDNA can be monitored in the recipient cells 24 to 72 hours after introduction (transient expression).

In addition, some vectors contain selectable markers such as the *gpt* (Mulligan and Berg, 1981, *Proc. Natl. Acad. Sci. USA* 78:2072-6) or *neo* (Southern and Berg, 1982, *J. Mol. Appl. Genet.* 1:327-41) bacterial genes. These selectable markers permit selection of transfected cells that exhibit stable, long-term expression of the vectors (and therefore the cDNA). The vectors can be maintained in the cells as episomal, freely replicating entities by using regulatory elements of viruses such as papilloma (Sarver et al., 1981, *Mol. Cell Biol.* 1:486) or Epstein-Barr (Sugden et al., 1985, *Mol. Cell Biol.* 5:410). Alternatively, one can also produce cell lines that have integrated the vector into genomic DNA. Both of these types of cell lines produce the gene product on a continuous basis. One can also produce cell lines that have amplified the number of copies of the vector (and therefore of the cDNA as well) to create cell lines that can produce high levels of the gene product (Alt et al., 1978, *J. Biol. Chem.* 253:1357).

The transfer of DNA into eukaryotic, in particular human or other mammalian cells, is now a conventional technique. The vectors are introduced into the recipient cells as pure DNA (transfection) by, for example, precipitation with calcium phosphate (Graham and vander Eb, 1973, *Virology* 52:466) or strontium phosphate (Brash et al., 1987, *Mol. Cell Biol.* 7:2013), electroporation (Neumann et al., 1982, *EMBO J.* 1:841), lipofection (Felgner et al., 1987, *Proc. Natl. Acad. Sci USA* 84:7413), DEAE dextran (McCuthan et al., 1968, *J. Natl. Cancer Inst.* 41:351), microinjection (Mueller et al., 1978, *Cell* 15:579), protoplast fusion (Schafner, 1980, *Proc. Natl. Acad. Sci. USA* 77:2163-7), or pellet guns (Klein et al., 1987, *Nature* 327:70). Alternatively, the cDNA can be introduced by infection with virus vectors. Systems are developed that use, for example, retroviruses (Bernstein et al., 1985, *Gen. Engrg.* 7:235), adenoviruses (Ahmad et al., 1986, *J. Virol.* 57:267), or Herpes virus (Spaete et al., 1982, *Cell* 30:295).

These eukaryotic expression systems can be used for studies of RRV genes and mutant forms of these genes, the RRV proteins and mutant forms of these proteins. Such uses include, for example, the identification of regulatory elements located in the 5' region of RRV genes on genomic clones that can be isolated from genomic DNA libraries, such as human or mouse libraries, using the information contained in the present invention. The eukaryotic expression systems may also be used to study the function of the normal complete protein, specific portions of the protein, or of naturally occurring or artificially produced mutant proteins. Naturally occurring RRV wild-type or mutant proteins may exist in a variety of cancers or diseases, while artificially

produced mutant proteins can be designed by site directed mutagenesis as described above. These latter studies may probe the function of any desired amino acid residue in the protein by mutating the nucleotide coding for that amino acid.

Using the above techniques, the expression vectors containing RRV genes or cDNA sequence or fragments or variants or mutants thereof can be introduced into human cells, mammalian cells from other species or non-mammalian cells as desired. The choice of cell is determined by the purpose of the treatment. For example, monkey COS cells (Gluzman, 1981, *Cell* 23:175-82) that produce high levels of the SV40 T antigen and permit the replication of vectors containing the SV40 origin of replication may be used. Similarly, Chinese hamster ovary (CHO), mouse NIH 3T3 fibroblasts or human fibroblasts or lymphoblasts may be used.

One method that can be used to express RRV polypeptides from the cloned RRV cDNA sequence in mammalian cells is to use the cloning vector, pXTI. This vector is commercially available from Stratagene, contains the Long Terminal Repeats (LTRs) and a portion of the GAG gene from Moloney Murine Leukemia Virus. The position of the viral LTRs allows highly efficient, stable transfection of the region within the LTRs. The vector also contains the Herpes Simplex Thymidine Kinase promoter (TK), active in embryonal cells and in a wide variety of tissues in mice, and a selectable neomycin gene conferring G418 resistance. Two unique restriction sites BglII and XhoI are directly downstream from the TK promoter. RRV cDNA, including the entire open reading frame for an RRV protein such as IL-6 and the 3' untranslated region of the cDNA is cloned into one of the two unique restriction sites downstream from the promoter.

The ligated product is transfected into mouse NIH 3T3 cells using Lipofectin (Life Technologies, Inc.) under conditions outlined in the product specification. Positive transfectants are selected after growing the transfected cells in 600 μg/ml G418 (Sigma, St. Louis, MO). The protein is released into the supernatant and may be purified by standard immunoaffinity chromatography techniques using antibodies raised against RRV proteins (see Example18).

Expression of RRV proteins in eukaryotic cells can be used as a source of proteins to raise antibodies. The RRV proteins may be extracted following release of the protein into the supernatant as described above, or, the cDNA sequence may be incorporated into a eukaryotic expression vector and expressed as a chimeric protein with, for example, β-globin. Antibody to β-globin is thereafter used to purify the chimeric protein. Corresponding protease cleavage sites engineered between the β-globin gene and the cDNA are then used to separate the two polypeptide fragments from one another after translation. One useful expression vector for generating β-globin chimeric proteins is pSG5 (Stratagene). This vector encodes rabbit β-globin.

The present invention thus encompasses recombinant vectors which comprise all or part of RRV genome or cDNA sequences, for expression in a suitable host. The RRV DNA is operatively linked in the vector to an expression control sequence in the recombinant DNA molecule so that a

RRV polypeptide can be expressed. The expression control sequence may be selected from the group consisting of sequences that control the expression of genes of prokaryotic or eukaryotic cells and their viruses and combinations thereof. The expression control sequence may be specifically selected from the group consisting of the lac system, the trp system, the tac system, the trc system, major operator and promoter regions of phage lambda, the control region of fd coat protein, the early and late promoters of SV40, promoters derived from polyoma, adenovirus, retrovirus, baculovirus and simian virus, the promoter for 3-phosphoglycerate kinase, the promoters of yeast acid phosphatase, the promoter of the yeast alpha-mating factors and combinations thereof.

The host cell, which may be transfected with the vector of this invention, may be selected from the group consisting of: *E. coli*, *Pseudomonas*, *Bacillus subtilis*, *Bacillus stearothermophilus* or other bacilli; other bacteria; yeast; fungi; plant; insect; mouse or other animal; or human tissue cells.

It is appreciated that for mutant or variant RRV DNA sequences, similar systems are employed to express and produce the mutant or variant product.

EXAMPLE 18

Production of Antibodies to RRV and RRV Proteins

Polyclonal or monoclonal antibodies (including humanized monoclonal antibodies) and fragments of monoclonal antibodies such as Fab, F(ab')2 and Fv fragments, as well as any other agent capable of specifically binding to an RRV protein, may be produced to the RRV virion or any of the RRV proteins (for example odd-numbered SEQ ID Nos 3-165). Optimally, antibodies raised against an RRV protein would specifically detect the RRV protein of interest (or a virion containing the protein of interest). That is, such antibodies would recognize and bind the protein and would not substantially recognize or bind to other proteins found in human or other cells. The determination that an antibody specifically detects the RRV protein is made by any one of a number of standard immunoassay methods; for instance, the Western blotting technique (Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

To determine that a given antibody preparation (such as one produced in a mouse) specifically detects the RRV protein by Western blotting, total cellular protein is extracted from murine myeloma cells and electrophoresed on a SDS-polyacrylamide gel. The proteins are then transferred to a membrane (for example, nitrocellulose) by Western blotting, and the antibody preparation is incubated with the membrane. After washing the membrane to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase; application of the substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium results in the production of

a dense blue compound by immuno-localized alkaline phosphatase. Antibodies which specifically detect an RRV protein will, by this technique, be shown to bind to the RRV protein band (which will be localized at a given position on the gel determined by its molecular weight). Non-specific binding of the antibody to other proteins (such as serum albumin) may occur and may be detectable as a weak signal on the Western blot. The non-specific nature of this binding will be recognized by one skilled in the art by the weak signal obtained on the Western blot relative to the strong primary signal arising from the specific antibody-VIAP protein binding.

A substantially pure virion can be obtained, or substantially pure RRV protein suitable for use as an immunogen is isolated by purification or recombinant expression. Concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms per milliliter. Monoclonal or polyclonal antibody to the protein can then be prepared as described by Harlow and Lane (Antibodies, A Laboratory Manual, Cold Spring Harbor Press. 1988).

Alternatively, antibodies may be raised against synthetic RRV peptides synthesized on a commercially available peptide synthesizer (see Example 26) based upon the predicted amino acid sequence of the RRV protein (Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Press. 1988).

Another method of raising antibodies against RRV proteins is by subcutaneous injection of a DNA vector which expresses the RRV protein into laboratory animals, such as mice. Delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford et al., 1987, *Particulate Sci. Technol.* 5:27-37) as described by Tang et al. (*Nature* 356:152-4, 1992). Expression vectors suitable for this purpose may include those which express the RRV protein under the transcriptional control of either the human β-actin promoter or the cytomegalovirus (CMV) promoter.

Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of the RRV protein identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler and Milstein (*Nature* 256:495, 1975) or derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall (*Enzymol.* 70:419, 1980), and derivative methods thereof.

Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow and Lane (Antibodies: A Laboratory Manual. 1988, Cold Spring Harbor Laboratory, New York).

Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein (for example see Example 17), which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis et al. (*J. Clin. Endocrinol. Metab.* 33:988-91, 1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony et al. (In: Handbook of Experimental Immunology, Wier, D. (ed.). Chapter 19. Blackwell. 1973). Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher (Manual of Clinical Immunology, Chapter 42. 1980).

Labeled Antibodies

Antibodies of the present invention can be conjugated with various labels for their direct detection (see Chapter 9, Harlow and Lane, Antibodies: A Laboratory Manual. 1988). The label, which may include, but is not limited to, a radiolabel, enzyme, fluorescent probe, or biotin, is chosen based on the method of detection available to the user.

EXAMPLE 19

Diagnostic Methods

An embodiment of the present invention is a method for screening a subject to determine if the subject has been infected with RRV. One major application of the RRV sequence information presented herein is in the area of diagnostic testing for predisposition to a disease (such as Kaposi's Sarcoma and lymphoproliferative disorders) that develops in at least a sub-set of hosts infected with RRV. The gene sequence of the RRV genes, including intron-exon boundaries

is also useful in such diagnostic methods. The method includes providing a biological sample obtained from the subject, in which sample includes DNA or RNA, and providing an assay for detecting in the biological sample the presence of any of the RRV genes or proteins. Suitable biological samples include samples obtained from body cells, such as those present in peripheral blood, urine, saliva, tissue biopsy, surgical specimen, fine needle aspirate specimen, amniocentesis samples and autopsy material. The detection in the biological sample may be performed by a number of methodologies, as outlined below.

The foregoing assay may be assembled in the form of a diagnostic kit and preferably comprises either: hybridization with oligonucleotides; PCR amplification of the gene or a part thereof using oligonucleotide primers; RT-PCR amplification of the RNA or a part thereof using oligonucleotide primers; or direct sequencing of any of the RRV genes present in a subject using oligonucleotide primers. The efficiency of these molecular genetic methods should permit the rapid identification of patients infected with RRV.

One embodiment of such detection techniques is the polymerase chain reaction amplification of reverse transcribed RNA (RT-PCR) of RNA isolated from cells (for example lymphocytes) followed by direct DNA sequence determination of the products. The presence of one or more RRV genes is taken as indicative of a potential RRV infection.

Alternatively, DNA extracted from lymphocytes or other cells may be used directly for amplification. The direct amplification from genomic DNA would be appropriate for analysis of an entire RRV gene including regulatory sequences located upstream and downstream from the open reading frame. Recent reviews of direct DNA diagnosis have been presented by Caskey (*Science* 236:1223-1228, 1989) and by Landegren et al. (*Science* 242:229-37, 1989).

Further studies of RRV genes isolated from subjects may reveal particular mutations, deletions or alterations in gene sequences, which occur at a high frequency within particular populations of individuals. In this case, rather than sequencing the entire RRV gene, it may be possible to design DNA diagnostic methods to specifically detect the most common RRV mutations, deletions or alterations in gene sequences.

The detection of specific DNA mutations or alterations in gene sequences may be achieved by methods such as hybridization using specific oligonucleotides (Wallace et al., 1986, *Cold Spring Harbor Symp. Quant. Biol.* 51:257-61), direct DNA sequencing (Church and Gilbert, 1984, *Proc. Natl. Acad. Sci. USA*. 81:1991-5), the use of restriction enzymes (Flavell et al., 1978, *Cell* 15:25; Geever et al., 1981, *Proc. Natl. Acad. Sci USA* 78:5081), discrimination on the basis of electrophoretic mobility in gels with denaturing reagent (Myers and Maniatis, 1986, *Cold Spring Harbor Symp. Quant. Biol.* 51:275-284), RNase protection (Myers et al., 1985, *Science* 230:1242), chemical cleavage (Cotton et al., 1985, *Proc. Natl. Acad. Sci. USA* 85:4397-401), and the ligase-mediated detection procedure (Landegren et al., 1988, *Science* 241:1077).

Oligonucleotides specific to normal, mutant or alterative sequences are chemically

synthesized using commercially available machines, labeled radioactively with isotopes (such as ^{32}P) or non-radioactively, with tags such as biotin (Ward and Langer et al., 1981. *Proc. Natl. Acad. Sci. USA* 78:6633-57), and hybridized to individual DNA samples immobilized on membranes or other solid supports by dot-blot or transfer from gels after electrophoresis. The presence of these specific sequences are visualized by methods such as autoradiography or fluorometric (Landegren et al., 1989, *Science* 242:229-37) or colorimetric reactions (Gebeyehu et al., 1987, *Nucleic Acids Res.* 15:4513-34). The absence of hybridization would indicate a mutation in the particular region of the gene, or that the patient is not infected with RRV.

Sequence differences between disclosed and other forms of RRV genes may also be revealed by the direct DNA sequencing method of Church and Gilbert (*Proc. Natl. Acad. Sci. USA* 81:1991-5, 1988). Cloned DNA segments may be used as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR (Wrichnik et al., 1987, *Nucleic Acids Res.* 15:529-42; Wong et al., 1987, *Nature* 330:384-6; Stoflet et al., 1988, *Science* 239:491-4). In this approach, a sequencing primer which lies within the amplified sequence is used with double-stranded PCR product or single-stranded template generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotides or by automatic sequencing procedures with fluorescent tags.

Sequence alterations may occasionally generate fortuitous restriction enzyme recognition sites or may eliminate existing restriction sites. Changes in restriction sites are revealed by the use of appropriate enzyme digestion followed by conventional gel-blot hybridization (Southern, 1975, *J. Mol. Biol.* 98:503). DNA fragments carrying the site (either normal, mutant, or alternative) are detected by their reduction in size or increase of corresponding restriction fragment numbers. Genomic DNA samples may also be amplified by PCR prior to treatment with the appropriate restriction enzyme; fragments of different sizes are then visualized under UV light in the presence of ethidium bromide after gel electrophoresis.

Screening based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing reagent. Small sequence deletions and insertions can be visualized by high-resolution gel electrophoresis. For example, a PCR product with small deletions is clearly distinguishable from a normal sequence on an 8% non-denaturing polyacrylamide gel (WO 91/10734; Nagamine et al., 1989, *Am. J. Hum. Genet.* 45:337-9). DNA fragments of different sequence compositions may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific "partial-melting" temperatures (Myers et al., 1985, *Science* 230:1242). Alternatively, a method of detecting a mutation comprising a single base substitution or other small change could be based on differential primer length in a PCR. For example, an invariant primer could be used in addition to a primer specific for a mutation. The PCR products of the normal and mutant genes can then be differentially

detected in acrylamide gels.

In addition to conventional gel-electrophoresis and blot-hybridization methods, DNA fragments may also be visualized by methods where the individual DNA samples are not immobilized on membranes. The probe and target sequences may be both in solution, or the probe sequence may be immobilized (Saiki et al., 1989, *Proc. Nat. Acad. Sci. USA* 86:6230-4). A variety of detection methods, such as autoradiography involving radioisotopes, direct detection of radioactive decay (in the presence or absence of scintillant), spectrophotometry involving calorigenic reactions and fluorometry involved fluorogenic reactions, may be used to identify specific individual genotypes.

If more than one mutation or alternative sequence is frequently encountered in one or more RRV genes, a system capable of detecting such multiple mutations would be desirable. For example, a PCR with multiple, specific oligonucleotide primers and hybridization probes may be used to identify all possible mutations or alternative sequences at the same time (Chamberlain et al., 1988, *Nucl. Acids Res.* 16:1141-55). The procedure may involve immobilized sequence-specific oligonucleotides probes (Saiki et al., 1989, *Proc. Nat. Acad. Sci. USA* 86:6230-4).

EXAMPLE 20

Quantitation of RRV Proteins

An alternative method of determining if a subject has been infected with RRV is to quantitate the level of one or more RRV proteins in the cells of a subject. This diagnostic tool would be useful for detecting the levels of RRV proteins which result from, for example, infection by RRV. These diagnostic methods, in addition to those described in EXAMPLE 19, provide an enhanced ability to diagnose susceptibility to diseases caused by RRV infection.

The determination of RRV protein levels would be an alternative or supplemental approach to the direct determination of the presence of one or more RRV genes by the methods outlined above in EXAMPLE 19. The availability of antibodies specific to one or more of the RRV proteins (for example those described in Example 18) will facilitate the quantitation of cellular RRV proteins by one of a number of immunoassay methods which are well known in the art and are presented in Harlow and Lane (Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, New York. 1988).

Such assays permit both the detection of RRV proteins in a biological sample and the quantitation of such proteins. Typical methods involve: providing a biological sample of the subject in which the sample contains cellular proteins, and providing an immunoassay for quantitating the level of RRV protein in the biological sample. This can be achieved by combining the biological sample with an RRV specific binding agent, such as an anti-RRV antibody (such as monoclonal or polyclonal antibodies), so that complexes form between the binding agent and the RRV protein present in the sample, and then detecting or quantitating such complexes.

In particular forms, these assays may be performed with the RRV specific binding agent immobilized on a support surface, such as in the wells of a microtiter plate or on a column. The biological sample is then introduced onto the support surface and allowed to interact with the specific binding agent so as to form complexes. Excess biological sample is then removed by washing, and the complexes are detected with a reagent, such as a second anti- RRV protein antibody that is conjugated with a detectable marker.

In an alternative assay, the cellular proteins are isolated and subjected to SDS-PAGE followed by Western blotting, for example as described in Example 18. After resolving the proteins, the proteins are transferred to a membrane, which is probed with specific binding agents that recognize any of the RRV proteins. The proteins are detected, for example with HRP-conjugated secondary antibodies, and quantitated.

In yet another assay, the level of one or more RRV proteins in cells is analyzed using microscopy. Using specific binding agents which recognize RRV, samples can be analyzed for the presence of one or more RRV proteins. For example, frozen biopsied tissue sections are thawed at room temperature and fixed with acetone at -200°C for 5 minutes. Slides are washed twice in cold PBS for 5 minutes each, then air-dried. Sections are covered with 20-30 μl of antibody solution (15-45 μg/ml) (diluted in PBS, 2% BSA at 15-50 μg/ml) and incubated at room temperature in humidified chamber for 30 minutes. Slides are washed three times with cold PBS 5 minutes each, allowed to air-dry briefly (5 minutes) before applying 20-30 μl of the second antibody solution (diluted in PBS, 2% BSA at 15-50 μg/ml) and incubated at room temperature in humidified chamber for 30 minutes. The label on the second antibody may contain a fluorescent probe, enzyme, radiolabel, biotin, or other detectable marker. The slides are washed three times with cold PBS 5 minutes each then quickly dipped in distilled water, air-dried, and mounted with PBS containing 30% glycerol. Slides can be stored at 4°C prior to viewing.

For samples prepared for electron microscopy (versus light microscopy), the second antibody is conjugated to gold particles. Tissue is fixed and embedded with epoxy plastics, then cut into very thin sections ($\sim$ 1-2 μm). The specimen is then applied to a metal grid, which is then incubated in the primary anti-RRV antibody, washed in a buffer containing BSA, then incubated in a secondary antibody conjugated to gold particles (usually 5-20 nm). These gold particles are visualized using electron microscopy methods.

For the purposes of quantitating the RRV proteins, a biological sample of the subject, which sample includes cellular proteins, is required. Such a biological sample may be obtained from body cells, such as those present in which expression of the protein has been detected. The expression of RRV proteins in peripheral blood leukocytes is clearly the most accessible and convenient source from which specimens can be obtained. Specimens can be obtained from peripheral blood, urine, saliva, tissue biopsy, amniocentesis samples, surgical specimens, fine needle aspirates, and autopsy material, particularly cancer cells. Quantitation of RRV proteins

would be made by immunoassay and compared to levels of the protein found in non-RRV expressing cells or to the level of RRV proteins in non-RRV infected cells (cells of the same origin that are not infected). A significant (preferably 50% or greater) increase in the amount of one or more RRV proteins in the cells of a subject compared to the amount of one or more RRV proteins found in non-RRV infected cells or that found in normal cells, would be taken as an indication that the subject may have been infected with RRV.

EXAMPLE 21

Sequence Variants of RRV

The amino acid sequence of the RRV proteins which are encoded by the RRV cDNAs (even-numbered SEQ ID NOS 2-164), are shown in odd-numbered SEQ ID NOS 3-165. Having presented the nucleotide sequence of the RRV genome and cDNAs and the amino acid sequence of these proteins, this invention now also facilitates the creation of DNA molecules, and thereby proteins, which are derived from those disclosed but which vary in their precise nucleotide or amino acid sequence from those disclosed. Such variants may be obtained through a combination of standard molecular biology laboratory techniques and the nucleotide sequence information disclosed by this invention.

Variant DNA molecules include those created by standard DNA mutagenesis techniques, for example, M13 primer mutagenesis. Details of these techniques are provided in Sambrook et al. (In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989, Chapter 15). By the use of such techniques, variants may be created which differ in minor ways from those disclosed. DNA molecules and nucleotide sequences which are derivatives of those specifically disclosed herein and which differ from those disclosed by the deletion, addition or substitution of nucleotides while still encoding a protein which possesses the functional characteristics of the RRV proteins are comprehended by this invention. Also within the scope of this invention are small DNA molecules which are derived from the disclosed DNA molecules. Such small DNA molecules include oligonucleotides suitable for use as hybridization probes or polymerase chain reaction (PCR) primers. As such, these small DNA molecules will comprise at least a segment of the RRV cDNA molecules or the RRV gene and, for the purposes of PCR, will comprise at least a 15 or a 20-50 nucleotide sequence of the RRV cDNAs (even-numbered SEQ ID Nos 2-164) or the RRV genes (i.e., at least 20-50 consecutive nucleotides of the RRV cDNA or gene sequences). DNA molecules and nucleotide sequences which are derived from the disclosed DNA molecules as described above may also be defined as DNA sequences which hybridize under stringent conditions to the DNA sequences disclosed, or fragments thereof.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing DNA used. Generally, the temperature of hybridization and the ionic strength (especially the Na^+ concentration) of the hybridization buffer will determine the stringency of hybridization. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed by Sambrook et al. (In: Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989 ch. 9 and 11), herein incorporated by reference. By way of illustration only, a hybridization experiment may be performed by hybridization of a DNA molecule (for example, a deviation of the RRV cDNA) to a target DNA molecule (for example, the RRV cDNA) which has been electrophoresed in an agarose gel and transferred to a nitrocellulose membrane by Southern blotting (Southern, *J. Mol. Biol.* 98:503, 1975), a technique well known in the art and described in Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989). Hybridization with a target probe labeled with [^{32}P]-dCTP is generally carried out in a solution of high ionic strength such as 6xSSC at a temperature that is 20-25°C below the melting temperature, T_m, described below. For such Southern hybridization experiments where the target DNA molecule on the Southern blot contains 10 ng of DNA or more, hybridization is typically carried out for 6-8 hours using 1-2 ng/ml radiolabeled probe (of specific activity equal to 10^9 CPM/μg or greater). Following hybridization, the nitrocellulose filter is washed to remove background hybridization. The washing conditions should be as stringent as possible to remove background hybridization but to retain a specific hybridization signal. The term T_m represents the temperature above which, under the prevailing ionic conditions, the radiolabeled probe molecule will not hybridize to its target DNA molecule. The T_m of such a hybrid molecule may be estimated from the following equation (Bolton and McCarthy, *Proc. Natl. Acad. Sci. USA* 48:1390, 1962): $T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - 0.63(\%\ formamide) - (600/l)$; where l = the length of the hybrid in base pairs.

This equation is valid for concentrations of Na^+ in the range of 0.01 M to 0.4 M, and it is less accurate for calculations of T_m in solutions of higher $[Na^+]$. The equation is also primarily valid for DNAs whose G+C content is in the range of 30% to 75%, and it applies to hybrids greater than 100 nucleotides in length (the behavior of oligonucleotide probes is described in detail in Ch. 11 of Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989).

Thus, by way of example, for a 150 base pair DNA probe derived from the open reading frame of the RRV cDNA (with a hypothetical %GC = 45%), a calculation of hybridization conditions required to give particular stringencies may be made as follows: For this example, it is assumed that the filter will be washed in 0.3 xSSC solution following hybridization, thereby: $[Na^+] = 0.045M$; %GC = 45%; Formamide concentration = 0; l = 150 base pairs; $T_m = 81.5 -$

$16.6(\log_{10}[Na^+]) + (0.41 \times 45) - (600/150)$; and so $T_m = 74.4°C$.

The T_m of double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology (Bonner et al., *J. Mol. Biol.* 81:123, 1973). Therefore, for this given example, washing the filter in 0.3 xSSC at 59.4-64.4°C will produce a stringency of hybridization equivalent to 90%; that is, DNA molecules with more than 10% sequence variation relative to the target RRV cDNA will not hybridize. Alternatively, washing the hybridized filter in 0.3 xSSC at a temperature of 65.4-68.4°C will yield a hybridization stringency of 94%; that is, DNA molecules with more than 6% sequence variation relative to the target RRV cDNA molecule will not hybridize. The above example is given entirely by way of theoretical illustration. One skilled in the art will appreciate that other hybridization techniques may be utilized and that variations in experimental conditions will necessitate alternative calculations for stringency.

In particular embodiments of the present invention, stringent conditions may be defined as those under which DNA molecules with more than 25%, 15%, 10%, 6% or 2% sequence variation (also termed "mismatch") will not hybridize.

The degeneracy of the genetic code further widens the scope of the present invention as it enables major variations in the nucleotide sequence of a DNA molecule while maintaining the amino acid sequence of the encoded protein. For example, the eleventh amino acid residue of the RRV MIP protein is alanine (SEQ ID NO 25). This is encoded in the RRV cDNA by the nucleotide codon triplet GCG. Because of the degeneracy of the genetic code, three other nucleotide codon triplets, GCT, GCA and GCC, also code for alanine. Thus, the nucleotide sequence of the RRV DNA could be changed at this position to any of these three codons without affecting the amino acid composition of the encoded protein or the characteristics of the protein. Based upon the degeneracy of the genetic code, variant DNA molecules may be derived from the DNA molecules disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of DNA sequences. DNA sequences which do not hybridize under stringent conditions to the DNA sequences disclosed by virtue of sequence variation based on the degeneracy of the genetic code are herein also comprehended by this invention.

The invention also includes DNA sequences that are substantially identical to any of the DNA sequences disclosed herein, where substantially identical means a sequence that has identical nucleotides in at least 75%, 80%, 85%, 90%, 95%, 98%, or even 99% of the aligned sequences.

One skilled in the art will recognize that the DNA mutagenesis techniques described above may be used not only to produce variant DNA molecules, but will also facilitate the production of proteins which differ in certain structural aspects from the RRV proteins, yet which proteins are clearly derivative of this protein and which maintain the essential characteristics of the RRV proteins. Newly derived proteins may also be selected in order to obtain variations on the characteristic of the RRV proteins, as described above. Such derivatives include those with variations in amino acid sequence including minor deletions, additions and substitutions.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random mutagenesis may be conducted at the target codon or region and the expressed protein variants screened for the optimal combination of desired activity. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence as described above are well known.

Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e., a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. Obviously, the mutations that are made in the DNA encoding the protein must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure.

Substitutional variants are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made conservatively, as defined above.

The effects of these amino acid substitutions or deletions or additions may be assessed for derivatives of the RRV proteins by assays in which DNA molecules encoding the derivative proteins are transfected into cells using routine procedures. These RRV proteins are expressed recombinantly (for example see Example 17), purified, and analyzed for their ability to cause symptoms associated with RRV infection, for example KS-like symptoms in rhesus macaque monkeys, as described in Examples 13 and 23.

EXAMPLE 22

Cloning RRV in Other Species

Having presented the genomic (SEQ ID NO 1) and cDNA nucleotide sequences of the rhesus macaque RRV (even-numbered SEQ ID Nos 2-164) and the amino acid sequence of the encoded proteins (odd-numbered SEQ ID Nos 3-165), this invention now also facilitates the identification of DNA molecules, and thereby proteins, which are the RRV homologs in other species. These other homologs can be derived from those sequences disclosed, but which vary in their precise nucleotide or amino acid sequence from those disclosed. Such variants may be obtained through a combination of standard molecular biology laboratory techniques and the nucleotide and amino acid sequence information disclosed by this invention.

The Japanese macaque RRV isolate was isolated from a lesion that was minced and co-cultured with primary rhesus fibroblasts. The isolate was then cloned by limiting dilution and a stock of virus generated from this clone. Total cellular DNA was harvested from virus infected cells and the DNA subjected to degenerate PCR for viral DNA polymerase, exactly as described

above for RRV. Once confirmed, a cosmid library of this virus was made from purified viral DNA (as described for RRV) and then a portion of the protein genes was cloned and sequenced.

Results for this analysis are shown in the following Table 1:

TABLE 1
RRV Sequences from Japanese Macaque

Total number of amino acid residues inferred: 972
Number of differences compared to RRV: 29
Percent identity: 97.02%

Japanese Macaque Data
These are end sequences. For ORFs represented twice, section I is from one plasmid, section II is from another plasmid. These are non-overlapping sections.

Orf 7 section I
GLFNSIDDTINALSRDCSVTFFQQANYTNVMRKQNELFTRLNSILCQGSAGSQKPATPSEPRT
ATVAATAASDVIKDAQYRKEQYMKKVARDGFKKLTECLQTQSAVLANALCMRVWGGVA
YGEASELVNHFLLRRRFVALPWEARCRSNQILFENSKYIKNSLYSQRLSREHVEIITLQFYGLI
TGPLTRQSDLFPGPANVVLAQCFEAAGMLPHHKMLVSEMIW

Orf 7 section II
PIESLFCGGLFNSIDDTINALSRDCSVTFFQQANYTNXMRKQNELFTRLNSILCQGSAGSXKP
ATPSEPRTATVXATAASDVIKDAQYRXEQYMKKVARDXFKKLTECLQTQSAVLANALCMR
RMGGRRI

Orf 8
YRKVATSVTVYRGWTETAVTGKQEVIRPVPQYEINHMDTTYQCFSSMRVNVNGIENTYTD
RDFTNQTVFLQPVEGLTDNIQRYFSQPVLYTTPGWFPGIYRVRTTVNCEIVDMIARSAEPYS
YFVTALGDTVEVSPFCLNDSTCSVADKAENGLGVRVLTNYTIVDFATRTPTTETRVFADSGE
YTVSWKAEDPKSAVCALTLWKTFPRAIQTTHESQLPLCGQRR

Orf 9 section I
VPSRFQTDIIPSGTVLKLLGRTENGTSVCVNVFRQQVYFYAKVPAGVNVTHVLQQALKNTA
GRAACGFSTRRVTKKILKTYDVAEHPVTEITLSSGSMLSTLSDRLVACGCEVFESNVDAVRR
FVLDHGFTTFGWYSCARATPRLAXRDARTALEFDCSWEDLSV

Orf 9 section II
MDFFNPYLGPRGPRPPSHKCTDAPAPAGAVQPPPDVCRLIPACLRTPGAGGMIPVTIPFPPTY
FENGARGDVLLAHERSMWTARGQRPVVPDPQDQSITFHAYDVVETTYAADRCAEV

Orf 10
AQMKIIYAPGDPNAEIVLGQSGPVLPTHTGGRVLGVYADAEKTIQPGSSAEVRVQLIFPTGSA
ARGDLAFLVAGVAPEPLFIVTPTLLLSGCTTHLRLFNPNGT

Orf 29b
NVAVEGNSSQDAGVAIATVLNEICSVPLSFLHHADKNTLIRSPIYMLGPEKAKAFESFIYALN
SGTFSASQTVVSHTIKLSFDPVAYLIDQIKAIRCIPLKDGGHTYCAKQKTMSDDVLVATVMA
HYMATNDKFVFKSLE

EXAMPLE 23

Screening Assays for Pharmaceutical Agents of Interest

The present disclosure provides a virus that is involved in the causation or progression of certain diseases, such as KS, and therefore provides an animal model and assays directed to identifying potential pharmaceutical agents that inhibit the biological activity of the virus. Drug screening assays which determine whether or not a drug has activity against the virus can include incubating a compound to be evaluated for use in treatment of the condition with cells which express the RRV proteins or peptides, and determining the effect of the compound on the activity of the virus. In vitro assays in which the virus is maintained in suitable cell culture are preferred, though in vivo animal models would also be effective.

In vitro assays include infecting cells such as rhesus fibroblasts, peripheral blood leukocytes or susceptible T cell lines such as MT-4 with the agent of interest in the presence of varying concentrations of compounds targeted against viral replication, including nucleoside analogs, chain terminators, antisense oligonucleotides and random polypeptides. (Asada et al., *J. Clin. Microbiol.* 27:2204, 1989; Kikuta et al., Lancet 7:861, 1989). Infected cultures and their supernatants can be assayed for the total amount of virus, including the presence of the viral genome, by quantitative PCR, by dot blot assays, or by using immunologic methods. For example, a culture of susceptible cells could be infected with the RRV in the presence of various concentrations of drug, fixed on slides after a period of days, and examined for viral antigen by indirect immunofluorescence with monoclonal antibodies to viral polypeptides (Kikuta et al, supra). Alternatively, chemically adhered MT-4 cell monolayers can be used for an infectious agent assay using indirect immunofluorescent antibody staining to search for focus reduction (Higashi, *J. Clin. Microbiol.* 27:2204, 1989, incorporated by reference).

As an alternative to whole cell in vitro assays, purified enzymes isolated from the RRV can be used as targets for rational drug design to determine the effect of the potential drug on enzyme activity, such as thymidylate sunthase or DNA polymerase. The genes for these two enzymes are provided herein. A measure of enzyme activity indicates an effect on the infectious agent itself. Drug screens using herpes viral products are known and have been previously described in EP 0514830 (herpes proteases) and WO 94/04920 (U_L 13 gene product).

In particular embodiments, this invention provides an assay for screening anti-KS chemotherapeutics. Infected cells can be incubated in the presence of a chemical agent that is a potential chemotherapeutic against KS (e.g. acyclo-guanosine). The level of virus in the cells is then determined after several days by IFA for antigens or Southern blotting for viral genome or Northern blotting for mRNA and compared to control cells. This assay can quickly screen large numbers of chemical compounds that may be useful against KS. This invention also provides an assay system that is employed to identify drugs or other molecules capable of binding to the DNA molecule or proteins, either in the cytoplasm or in the nucleus, thereby inhibiting or potentiating transcriptional activity. This assay would be useful in the development of drugs that

are specific against particular cellular activity, or that would potentiate such activity, in time or in level of activity. Also included are drugs identified by this assay which have an anti-viral activity, and an effect against conditions associated with RRV infection, such as KS.

EXAMPLE 24

Generating Animal Models

Animal models are useful for resolving a number of fundamental problems of infectious diseases that include, but are not limited to, determinants of virulence of the organism, mechanisms of host resistance, mechanisms of pathogenicity, establishment and regulation of chronic infection, and antimicrobial and chemotherapeutic actions of drugs on infectious agents. Variables that are commonly manipulated to address fundamental problems include, but are not limited to, the strain of infectious agent, the infecting dose of infectious agent and the route of administration of the infectious agent, the species or subspecies of animal, the age of animal, and the genetic background of the animal (Viral pathogenesis, N. Nathanson, Lippincot-Raven, Philadelphia, 1997).

In an embodiment in which one or more RRV strains are employed for generating an animal model, the RRV used may be naturally occurring variant isolates recovered from rhesus macaques and other non-human primate species, molecular clones generated from these naturally occurring variant isolates and recombinant viruses with introduced mutations, deletions or recombined genomes designed to address function of specific genes.

By manipulating the infecting dose and route of RRV administration virus-host interactions dependent upon dose and tissue or organ-specific disease manifestations can be explored. Thus, the present invention includes various doses of RRV administered by oral, inhalation, intratracheal, intravaginal, intrarectal and parenteral routes including, but not limited to intravenous, intraarterial, intradermal, subcutaneous, intramuscular, intraperitoneal and organ-specific administration routes such and intracerebral and intraocular administration.

Many disease manifestations with a given infections agent are highly influenced by age and species or subspecies of the host and the particular genetic makeup of the host. The present disclosure provides a virus that is involved in the causation or progression of certain diseases, such as KS, in the rhesus macaque, but is also useful for the study of and discovery of disease manifestations that are host species, age and genetic background dependent. In particular embodiments, one skilled in the art may vary the species of animal to which the RRV is administered to produce or discover a particular disease manifestation, or similarly vary the genetic background of the animal to produce or discover a particular disease manifestation, even including the use of genetically engineered animals.

EXAMPLE 25

Pharmaceutical Compositions and Modes of Administration

Various delivery systems for administering pharmaceutical proteins from the RRV include encapsulation in liposomes, microparticles, microcapsules, expression by recombinant cells, receptor-mediated endocytosis (see Wu and Wu, *J. Biol. Chem.* 1987, 262:4429-32), and construction of a therapeutic nucleic acid (such as an anti-sense molecule) as part of a retroviral or other vector. Methods of introduction include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, the pharmaceutical compositions may be introduced into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir.

The use of liposomes as a delivery vehicle is another delivery method of the present invention. The liposomes fuse with the target site and deliver the contents of the lumen intracellularly. The liposomes are maintained in contact with the target cells for a sufficient time for fusion to occur, using various means to maintain contact, such as isolation and binding agents. Liposomes may be prepared with purified proteins or peptides that mediate fusion of membranes, such as Sendai virus or influenza virus. The lipids may be any useful combination of known liposome forming lipids, including cationic lipids, such as phosphatidylcholine. Other potential lipids include neutral lipids, such as cholesterol, phosphatidyl serine, phosphatidyl glycerol, and the like. For preparing the liposomes, the procedure described by Kato et al. (*J. Biol. Chem.* 1991, 266:3361) may be used.

The present invention also provides pharmaceutical compositions which include a therapeutically effective amount of one or more RRV proteins or DNA, alone or with a pharmaceutially acceptable carrier.

The pharmaceutical compositions or methods of treatment may be administered in combination with other therapeutic treatments, such as other antineoplastic or antitumorigenic therapies.

Administration of Nucleic Acid Molecules

In an embodiment in which one or more RRV nucleic acids are employed for generating an animal model, the analog may be delivered intracellularly (e.g., by expression from a nucleic acid vector or by receptor-mediated mechanisms). In a specific embodiment where the therapeutic molecule is a nucleic acid, administration may be achieved by an appropriate nucleic acid expression vector which is administered so that it becomes intracellular, e.g., by use of a retroviral

vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see e.g., Joliot et al., *Proc. Natl. Acad. Sci. USA* 1991, 88:1864-8). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The vector pCDNA, is an example of a method of introducing the foreign cDNA into a cell under the control of a strong viral promoter (CMV) to drive the expression. However, other vectors can be used. Other retroviral vectors (such as pRETRO-ON, Clontech), also use this promoter but have the advantages of entering cells without any transfection aid, integrating into the genome of target cells ONLY when the target cell is dividing (as cancer cells do, especially during first remissions after chemotherapy) and they are regulated. It is also possible to turn on the expression of the RRV nucleic acid by administering tetracycline when these plasmids are used. Hence these plasmids can be allowed to transfect the cells, then administer a course of tetracycline with a course of chemotherapy to achieve better cytotoxicity.

Other plasmid vectors, such as pMAM-neo (also from Clontech) or pMSG (Pharmacia) use the MMTV-LTR promoter (which can be regulated with steroids) or the SV10 late promoter (pSVL, Pharmacia) or metallothionein - responsive promoter (pBPV, Pharmacia) and other viral vectors, including retroviruses. Examples of other viral vectors include adenovirus, AAV (adeno-associated virus), recombinant HSV, poxviruses (vaccinia) and recombinant lentivirus (such as HIV). All these vectors achieve the basic goal of delivering into the target cell the cDNA sequence and control elements needed for transcription. The present invention includes all forms of nucleic acid delivery, including synthetic oligos, naked DNA, plasmid and viral, integrated into the genome or not.

Also contemplated are inhibitory nucleic acid therapeutics which can inhibit the activity of RRV, for example in subject with KS or other diseases associated with RRV infection. Inhibitory nucleic acids may be single-stranded nucleic acids, which can specifically bind to a complementary nucleic acid sequence. By binding to the appropriate target sequence, an RNA-RNA, a DNA-DNA, or RNA-DNA duplex or triplex is formed. These nucleic acids are often termed "antisense" because they are usually complementary to the sense or coding strand of the gene, although recently approaches for use of "sense" nucleic acids have also been developed. The term "inhibitory nucleic acids" as used herein, refers to both "sense" and "antisense" nucleic acids.

By binding to the target nucleic acid, the inhibitory nucleic acid can inhibit the function of the target nucleic acid. This could, for example, be a result of blocking DNA transcription, processing or poly(A) addition to mRNA, DNA replication, translation, or promoting inhibitory mechanisms of the cells, such as promoting RNA degradation. Inhibitory nucleic acid methods therefore encompass a number of different approaches to altering expression of RRV genes.

Cleavage, and therefore inactivation, of the target nucleic acids may be effected by attaching a substituent to the inhibitory nucleic acid which can be activated to induce cleavage reactions. The substituent can be one that affects either chemical, or enzymatic cleavage. Alternatively, cleavage can be induced by the use of ribozymes or catalytic RNA. In this approach, the inhibitory nucleic acids would include either naturally occurring RNA (ribozymes) or synthetic nucleic acids with catalytic activity.

The inhibitory nucleic acid therapies can be used to target nucleic acids to sequences of RRV for use in treating conditions caused by the RRV, or proteins of the RRV, for example for treating KS or KS-like syndromes.

Administration of Antibodies

Therapeutic, intravenous, polyclonal or monoclonal antibodies has been used as a mode of passive immunotherapy of herpesviral diseases, such as infection with CMV. Immune globulin from subjects previously infected with the RRV and bearing a suitably high titer of antibodies against the virus can be given in combination with antiviral agents (e.g. ganciclovir), or in combination with other modes of immunotherapy that are currently being evaluated for the treatment of KS, which are targeted to modulating the immune response (i.e. treatment with copolymer-1, antiidiotypic monoclonal antibodies, T cell "vaccination"). Antibodies specific for an epitope expressed on cells infected with the RRV are preferred and can be obtained as described above.

The present invention also provides pharmaceutical compositions which include a therapeutically effective amount of the antibody, and a pharmaceutically acceptable carrier or excipient.

EXAMPLE 26

Vaccines

This invention provides substances suitable for use as vaccines for the prevention of diseases associated with RRV infection, such as KS, and methods for administering them. The vaccines are directed against RRV, and may include antigens obtained from RRV. In one embodiment, the vaccine contains attenuated RRV. In another embodiment, the vaccine contains killed RRV. In another embodiment, the vaccine contains a nucleic acid vector encoding RRV, or a surface protein, such as a capsid protein. In another embodiment, the vaccine is a subunit vaccine containing an RRV subunit, such as glycoprotein B, major capsid protein, or other gene products found to elicit appropriate humoral and/or cell mediated immune responses.

This invention also provides a method of vaccinating a subject against Kaposi's sarcoma and lymphoproliferative disorders, comprising administering to a susceptible subject an effective amount of the peptide or polypeptide encoded by an isolated DNA molecule encoding a

polypeptide or combination of polypeptides expressed by the DNA molecule, and a suitable acceptable carrier. In one embodiment, naked DNA is administered to the subject in an effective amount to vaccinate the subject against Kaposi's sarcoma and lymphoproliferative disorders, or other disease associated with RRV infection.

The vaccine can be made using synthetic peptide or recombinantly-produced polypeptide described above as antigen. Typically, a vaccine will include from about 1 to 50 micrograms of antigen, for example from about 15 to about 45 micrograms. Typically, the vaccine is formulated so that a dose includes about 0.5 milliliters. The vaccine may be administered by any route known in the art, for example parenteral, subcutaneous or intramuscular.

There are a number of strategies for amplifying an antigen's effectiveness, particularly as related to the art of vaccines. For example, cyclization of a peptide can increase the peptide's antigenic and immunogenic potency. See U.S. Pat. No. 5,001,049. More conventionally, an antigen can be conjugated to a suitable carrier, usually a protein molecule. This procedure can allow multiple copies of an antigen, such as a peptide, to be conjugated to a single larger carrier molecule. Additionally, the carrier may possess properties which facilitate transport, binding, absorption or transfer of the antigen.

For parenteral administration, such as subcutaneous injection, examples of suitable carriers are the tetanus toxoid, the diphtheria toxoid, serum albumin and lamprey, or keyhole limpet, hemocyanin because they provide the resultant conjugate with minimum genetic restriction. Conjugates including these universal carriers can function as T cell clone activators in individuals having very different gene sets. The conjugation between a peptide and a carrier can be accomplished using one of the methods known in the art. Specifically, the conjugation can use bifunctional cross-linkers as binding agents as detailed, for example, by Means and Feeney, "A recent review of protein modification techniques," *Bioconjugate Chem.* 1:2-12 (1990).

Vaccines against RRV can be made from the RRV envelope glycoproteins. These proteins can be purified and used for vaccination (Lasky, L. A., 1990, *J. Med. Virol.* 31:59). MHC-binding peptides from cells infected with the human herpesvirus can be identified for vaccine candidates per the methodology of Marloes, et al., 1991, *Eur. J. Immunol.* 21:2963-2970. The RRV antigen may be combined or mixed with various solutions and other compounds as is known in the art. For example, it may be administered in water, saline or buffered vehicles with or without various adjuvants or immunodiluting agents. Examples of such adjuvants or agents include aluminum hydroxide, aluminum phosphate, aluminum potassium sulfate (alum), beryllium sulfate, silica, kaolin, carbon, water-in-oil emulsions, oil-in-water emulsions, muramyl dipeptide, bacterial endotoxin, lipid X, Corynebacterium parvum (Propionibacterium acnes), Bordetella pertussis, polyribonucleotides, sodium alginate, lanolin, lysolecithin, vitamin A, saponin, liposomes, levamisole, DEAE-dextran, blocked copolymers or other synthetic adjuvants. Such adjuvants are available commercially from various sources, for example, Merck Adjuvant 65

(Merck and Company, Inc., Rahway, N.J.) or Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, Mich.). Other suitable adjuvants are Amphigen (oil-in-water), Alhydrogel (aluminum hydroxide), or a mixture of Amphigen and Alhydrogel. Only aluminum is approved for human use.

The proportion of antigen and adjuvant can be varied over a broad range so long as both are present in effective amounts. For example, aluminum hydroxide can be present in an amount of about 0.5% of the vaccine mixture (Al_2O_3 basis). On a per-dose basis, the amount of the antigen can range from about 0.1 μg to about 100 μg protein per subject, for example about 1 μg to about 50 μg per dose, or about 15 μg to about 45 μg. A suitable dose size is about 0.5 ml. Accordingly, a dose for intramuscular injection, for example, would comprise 0.5 ml containing 45 μg of antigen in admixture with 0.5% aluminum hydroxide. After formulation, the vaccine may be incorporated into a sterile container which is then sealed and stored at a low temperature, for example 4°C., or it may be freeze-dried. Lyophilization permits long-term storage in a stabilized form.

The vaccines may be administered by any conventional method for the administration of vaccines including oral and parenteral (e.g., subcutaneous or intramuscular) injection. Intramuscular administration is preferred. The treatment may consist of a single dose of vaccine or a plurality of doses over a period of time. Also, the antigen could be a component of a recombinant vaccine which could be adaptable for oral administration. Vaccines of the invention may be combined with other vaccines for other diseases to produce multivalent vaccines. A pharmaceutically effective amount of the antigen can be employed with a pharmaceutically acceptable carrier such as a protein or diluent useful for the vaccination of mammals, particularly humans. Other vaccines may be prepared according to methods well-known to those skilled in the art.

Those of skill will readily recognize that it is only necessary to expose a mammal to appropriate epitopes in order to elicit effective immunoprotection. The epitopes are typically segments of amino acids which are a small portion of the whole protein. Using recombinant genetics, it is routine to alter a natural protein's primary structure to create derivatives embracing epitopes that are identical to or substantially the same as (immunologically equivalent to) the naturally occurring epitopes. Such derivatives may include peptide fragments, amino acid substitutions, amino acid deletions and amino acid additions of the amino acid sequence for the viral polypeptides from the human herpesvirus. For example, it is known in the protein art that certain amino acid residues can be substituted with amino acids of similar size and polarity without an undue effect upon the biological activity of the protein. The human herpesvirus proteins have significant tertiary structure and the epitopes are usually conformational. Thus, modifications should generally preserve conformation to produce a protective immune response.

EXAMPLE 27

Peptide Synthesis and Purification

The peptides provided by the present invention can be chemically synthesized by any of a number of manual or automated methods of synthesis known in the art. For example, solid phase peptide synthesis (SPPS) is carried out on a 0.25 millimole (mmole) scale using an Applied Biosystems Model 431A Peptide Synthesizer and using 9-fluorenylmethyloxycarbonyl (Fmoc) amino-terminus protection, coupling with dicyclohexylcarbodiimide/ hydroxybenzotriazole or 2-(1H-benzo-triazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate/ hydroxybenzotriazole (HBTU/HOBT), and using p-hydroxymethylphenoxymethylpolystyrene (HMP) or Sasrin resin for carboxyl-terminus acids or Rink amide resin for carboxyl-terminus amides.

Fmoc-derivatized amino acids are prepared from the appropriate precursor amino acids by tritylation and triphenylmethanol in trifluoroacetic acid, followed by Fmoc derivitization as described by Atherton et al. (Solid Phase Peptide Synthesis, IRL Press: Oxford, 1989).

Sasrin resin-bound peptides are cleaved using a solution of 1% TFA in dichloromethane to yield the protected peptide. Where appropriate, protected peptide precursors are cyclized between the amino- and carboxyl-termini by reaction of the amino-terminal free amine and carboxyl-terminal free acid using diphenylphosphorylazide in nascent peptides wherein the amino acid sidechains are protected.

HMP or Rink amide resin-bound products are routinely cleaved and protected sidechain-containing cyclized peptides deprotected using a solution comprised of trifluoroacetic acid (TFA), optionally also comprising water, thioanisole, and ethanedithiol, in ratios of 100 : 5 : 5 : 2.5, for 0.5 - 3 hours at room temperature.

Crude peptides are purified by preparative high pressure liquid chromatography (HPLC), for example using a Waters Delta-Pak C18 column and gradient elution with 0.1% TFA in water modified with acetonitrile. After column elution, acetonitrile is evaporated from the eluted fractions, which are then lyophilized. The identity of each product so produced and purified may be confirmed by fast atom bombardment mass spectroscopy (FABMS) or electrospray mass spectroscopy (ESMS).

Having illustrated and described the principles of cloning the RRV genome, cDNA, proteins encoded by the cDNA, and modes of use of these biological molecules, it should be apparent to one skilled in the art that the invention can be modified in arrangement and detail without departing from such principles. In view of the many possible embodiments to which the principles of our invention may be applied, it should be recognized that the illustrated embodiments are only examples of the invention and should not be taken as a limitation on the scope of the invention. Rather, the scope of the invention is in accord with the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

CLAIMS AS AMENDED UNDER ARTICLE 34

1. An isolated virus (RRV) as deposited with ATCC as deposit accession number VR-2601.

2. A purified virus, having a nucleic acid sequence
   (a) shown in SEQ ID NO 1 or
   (b) a conservative variant thereof.

3. The purified virus of claim 2, wherein the nucleic acid sequence has at least 95% sequence identity to the nucleic acid sequence shown in SEQ ID NO 1.

A

4. A purified protein encoded by an open reading frame of the virus of claim 2.

5. A purified protein of claim 4, wherein the protein comprises an amino acid sequence selected from the group consisting of:
   (a) an amino acid sequence shown in odd numbered sequences of SEQ ID NOS. 3-165; and
   (b) amino acid sequences that differ from those specified in (a) by one or more conservative amino acid substitutions wherein the function of the protein is preserved.

6. A purified protein with an amino acid sequence that is at least 95% sequence identity to the sequences specified in claim 5(a) or 5(b).

7. The purified protein of claim 5, wherein the amino acid sequence is selected from odd numbered sequences within the group consisting of SEQ ID NOS 3-19 and 23-165.

8. An isolated nucleic acid molecule encoding a protein according to claim 5.

9. An isolated nucleic acid molecule according to claim 8, wherein the molecule comprises a sequence selected from the group consisting of even numbered sequences of SEQ ID NOS 2-164.

10. The isolated nucleic acid molecule according to claim 9, wherein the molecule comprises a sequence selected from the group consisting of even numbered sequences of SEQ ID NOS 2-18 and 22-164.

11. A recombinant nucleic acid molecule comprising a promoter sequence operably linked to a nucleic acid molecule according to claim 8.

12. A cell transformed with a recombinant nucleic acid molecule according to claim 8.

13. A non-human mammal purposefully infected with the virus of claim 2.

14. The mammal of claim 13, wherein the mammal is a primate.

15. An oligonucleotide comprising a sequence selected from the group consisting of:

A

(a) at least 20 contiguous nucleotides of the nucleic acid sequence of the virus of claim 2;

(b) at least 30 contiguous nucleotides of the nucleic acid sequence of the virus of claim 2; and

(c) at least 50 contiguous nucleotides of the nucleic acid sequence of the virus of claim 2.

16. An isolated nucleic acid molecule that:

(a) hybridizes under stringent conditions with a nucleic acid probe comprising the sequence of claim 15; and

(b) encodes a protein of claim 6.

17. An isolated nucleic acid molecule encoding a protein of claim 6.

18. An isolated nucleic acid molecule encoding all proteins encoded by the virus of claim 2, and having a biological activity of an RRV virus.

19. A method for testing the efficacy of a drug in the treatment of a condition associated with the virus of claim 2, the method comprising:

(a) administering the drug to a non-human primate infected with the virus of claim 2; and

(b) observing the primate to determine if the drug prevents or reduces the presentation of one or more symptoms associated with viral infection.

20. The method of claim 19, wherein the primate is immunocompromised.

21. The method of claim 20, wherein the drug is for the treatment of Kaposi's sarcoma and lymphoproliferative disorders.

22. The method of claim 20, wherein the primate is immuno-compromised as a result of infection by Simian Immunodeficiency Virus (SIV).

23. The method of claim 19, wherein the condition associated with infection with the virus of claim 2 is one or more of B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobinulinemia or autoimmune hemolytic anemia.

24. The method of claim 19, wherein the non-human primate is a Rhesus macaque monkey.

A

25. A method for producing a non-human primate model for testing potential treatments for a condition associated an infection with the virus of claim 2, comprising

(a) administering a treatment to the primate to render the primate immunocompromised; and

(b) infecting the primate with the virus of claim 2.

26. The method of claim 25, wherein the condition is Kaposi's sarcoma and lymphoproliferative disorders.

27. The method of claim 25 wherein the treatment used to render the primate immuno-compromised is infection with SIV.

28. The method of claim 25 wherein the non-human primate is a Rhesus macaque monkey.

29. A method for testing the efficacy of a candidate vaccine against the virus of claim 2, or conditions associated infection with virus of claim 2, the method comprising:

(a) administering the vaccine to a subject capable of infection with the virus of claim 2;

(b) inoculating the subject with the virus; and

(c) observing the subject to determine if the vaccine prevents or reduces an incidence of viral infection or presentation of one or more conditions associated with the viral infection.

30. The method of claim 29, wherein the subject is a primate.

A

31. The method of claim 30, wherein the primate is a non-human primate.

32. The method of claim 29, wherein the primate is immunocompromised.

33. The method of claim 29, wherein the conditions associated with infection include B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobinulinemia or autoimmune hemolytic anemia.

34. The method of claim 31, wherein the non-human primate is a Rhesus macaque monkey.

add A1

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/38, 7/00, 5/10, 15/00, A01K 67/027, C07K 14/03 — A2

(11) International Publication Number: WO 00/28040

(43) International Publication Date: 18 May 2000 (18.05.00)

(21) International Application Number: PCT/US99/26260

(22) International Filing Date: 5 November 1999 (05.11.99)

(30) Priority Data:

60/107,507	6 November 1998 (06.11.98)	US
60/109,409	20 November 1998 (20.11.98)	US

(71) Applicant *(for all designated States except US)*: OREGON HEALTH SCIENCES UNIVERSITY [US/US]; Office of Technology Management, 3181 S.W. Sam Jackson Park Road, L335, Portland, OR 97201-3098 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: WONG, Scott, W. [US/US]; 7926 S.W. 167th Place, Beaverton, OR 97007 (US). AXTHELM, Michael, K. [US/US]; 1345 S.E. 35th Avenue, Hillsboro, OR 97123 (US). SEARLES, Robert, P. [US/US]; 18300 N.W. Heritage Parkway #106, Beaverton, OR 97006 (US).

(74) Agent: NOONAN, William, D.; Klarquist, Sparkman, Campbell, Leigh & Whinston, LLP, Suite 1600 – One World Trade Center, 121 S.W. Salmon Street, Portland, OR 97204 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

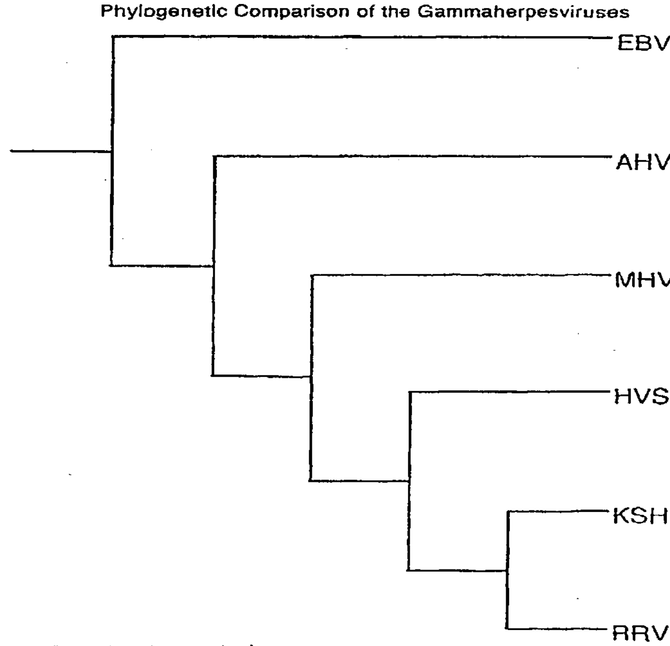
Without international search report and to be republished upon receipt of that report.

(54) Title: CLONING OF RHADINOVIRUS GENOME AND METHODS FOR ITS USE

(57) Abstract

A novel rhesus macaque rhadinovirus, herein designated RRV, is disclosed. The genomic, cDNA and proteins sequences are provided. RRV has some similarity to human Kaposi's sarcoma–associated herpesvirus and causes Kaposi's sarcoma–like symptoms in immuno–compromised non–human primates. RRV possesses genes for both Interleukin–6 and macrophage inflammatory protein 1. The genome of RRV is useful for research, clinical and diagnostic applications aimed towards the rhadinoviruses and herpesviruses in general and KSHV in particular. In addition, methods for using RRV to produce a non–human primate model for the testing of Kaposi's sarcoma–associated herpesvirus therapeutics and vaccines are presented.

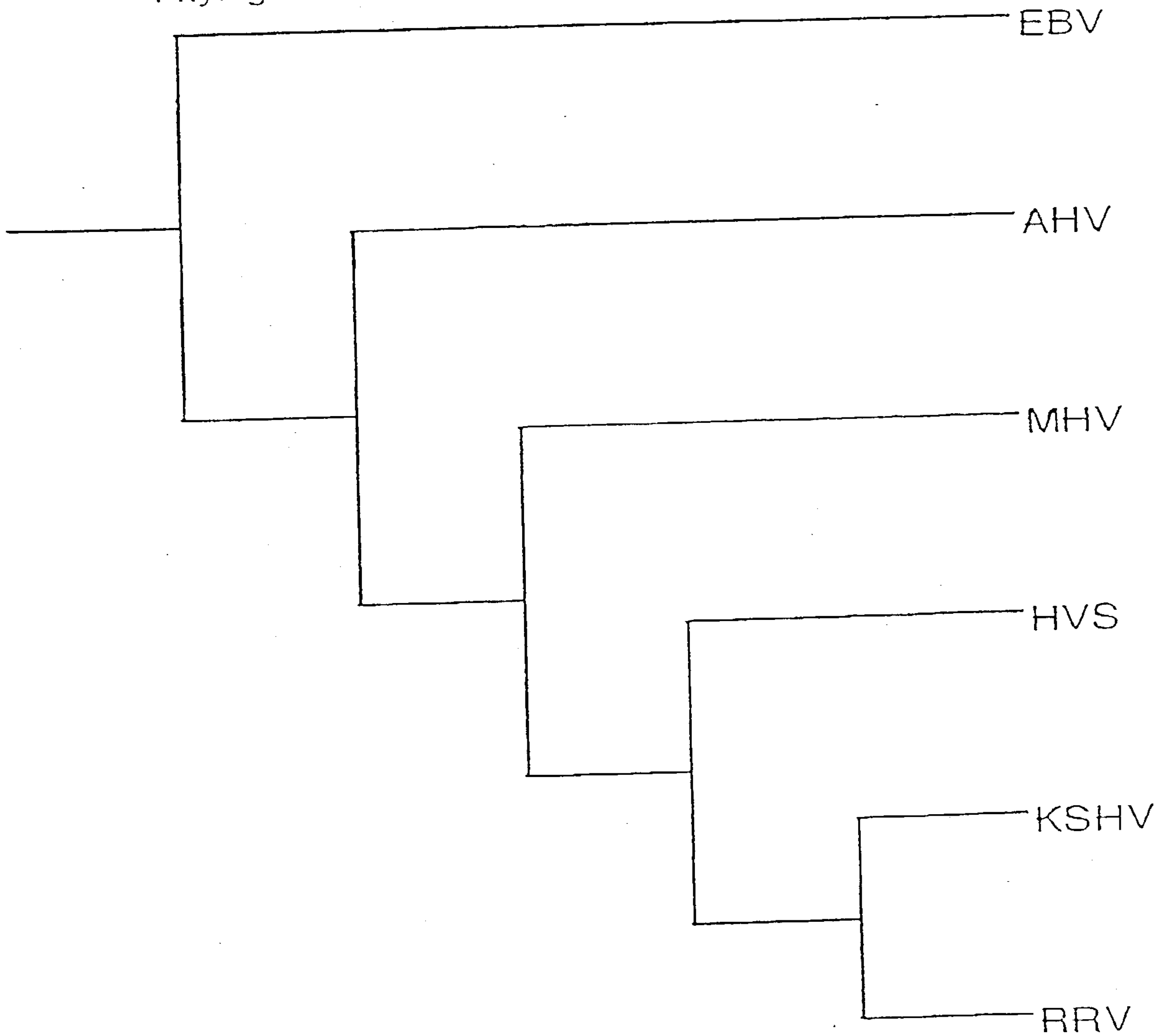
Phylogenetic Comparison of the Gammaherpesviruses



EBV - Epstein-Barr virus
AHV - Alcelaphine herpesvirus
MHV - Murine herpesvirus 68
HVS - Herpesvirus saimiri
KSHV - Kaposi's sarcoma-associated herpesvirus
RRV - Rhesus rhadinovirus 17577

FIG. 1

Phylogenetic Comparison of the Gammaherpesviruses



EBV - Epstein-Barr virus
AHV - Alcelaphine herpesvirus
MHV - Murine herpesvirus 68
HVS - Herpesvirus saimiri
KSHV - Kaposi's sarcoma-associated herpesvirus
RRV - Rhesus rhadinovirus 17577

FIG. 2

Restriction Fragments of the RRV 17577 Genome

BamHI		EcoRI		Hind III	
fragment number	fragment size (bp)	fragment number	fragment size (bp)	fragment number	fragment size (bp)
1	17189	1	12476	1*	22006
2	15598	2	10342	2	17108
3	15441	3	9565	3	16542
4	12360	4	9213	4	14134
5*	8943	5	8465	5	11516
6	7747	6	8036	6	10743
7	7718	7	7969	7	8452
8	7142	8	7416	8	5995
9	6667	9	7278	9	4679
10	6474	10	7002	10	3374
11	6333	11*	5400	11	2963
12	3978	12*	5054	12	2891
13	3411	13	4907	13	2849
14*	3157	14	4771	14*	2832
15	3008	15	4272	15	1599
16	2916	16	4099	16	1272
17	2210	17	3516	17	1016
18	1343	18	2102	18	853
		19	1868	19	811
		20	1603		
		21	1512		
		22	1221		
		23	910		
		24	624		
		25	609		
		26	592		
		27	584		
		28	122		
		29	107		

* Indicates that the fragment size excludes terminal repeat sequences

FIG. 3

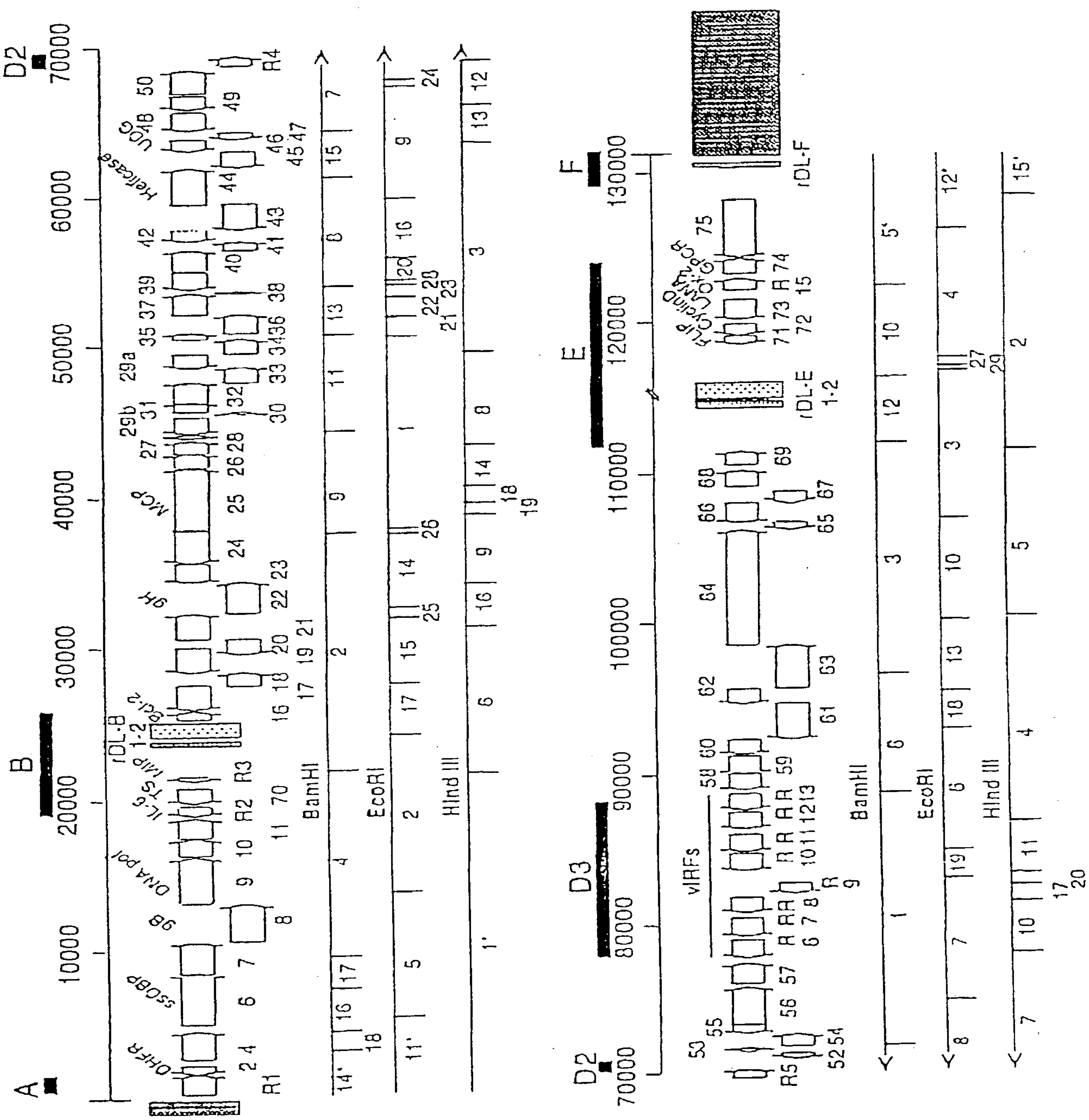

A
B
D2
D3
E
F
DHFR
ssDBP
9B
DNA pol
IL-6
vIRFs
Helicase
BamHI
EcoRI
HindIII
rDL-B 1-2
rDL-E 1-2
rDL-F

FIG. 4

MacVector Output for long unique region of rhesus rhadinovirus 17577

```
LOCUS       LONG UNIQU 131634 BP DS-DNA          UPDATED   06/26/98
DEFINITION  -
ACCESSION   -
KEYWORDS    -
SOURCE      -
FEATURES                    Description
    pept                 1 R1
    pept               (C1 Similar to HHV8 Orf 2 - dihydrofolate reductase
    pept                 1 Similar to HHV8 Orf 4 - complement binding
                           protein
    pept                 1 Similar to HHV8 Orf 6 - ssDNA binding protein
    pept                 1 Similar to HHV8 Orf 7 - transport protein
    pept                 1 Similar to HHV8 Orf 8 - glycoprotein B
    pept                 1 Similar to HHV8 Orf 9 - DNA polymerase
    pept                 1 Similar to HHV8 Orf 10
    pept                 1 Similar to HHV8 Orf 11
    pept               (C1 R2 viral IL-6
    pept               (C1 Similar to HHV8 Orf 70
    pept               (C1 R3 similar to HHV8 MIP
    frag               (C) misc. feature MIP homology, but no initiation
                           codon
    rpt                    repeat sequence
    rpt                    repeat sequence
    pept                 1 Similar to HHV8 Orf 16 - Bcl-2 homolog
    pept               (C1 Similar to HHV8 Orf 17 - capsid protein
    pept                 1 Similar to HHV8 Orf 18
    pept               (C1 Similar to HHV8 Orf 19 - tegument protein
    pept               (C1 Similar to HHV8 Orf 20
    pept                 1 Similar to HHV8 Orf 21 - thymidine kinase
    pept                 1 Similar to HHV8 Orf 22 - glycoprotein H
    pept               (C1 Similar to HHV8 Orf 23
    pept               (C1 Similar to HHV8 Orf 24
    pept                 1 Similar to HHV8 Orf 25 - major capsid protein
    pept                 1 Similar to HHV8 Orf 26 - capsid protein
    pept                 1 Similar to HHV8 Orf 27
    pept                 1 Similar to HHV8 Orf 28
    pept               (C1 Similar to HHV8 Orf 29b
    pept                 1 Similar to HHV8 Orf 30
    pept                 1 Similar to HHV8 Orf 31
    pept                 1 Similar to HHV8 Orf 32
    pept                 1 Similar to HHV8 Orf 33
    pept               (C1 Similar to HHV8 Orf 29a
    pept                 1 Similar to HHV8 Orf 34
    pept                 1 Similar to HHV8 Orf 35
    pept                 1 Similar to HHV8 Orf 36 -kinase
    pept                 1 Similar to HHV8 Orf 37 - alkaline exonuclease
    pept                 1 Similar to HHV8 Orf 38
    pept               (C1 Similar to HHV8 Orf 39 - glycoprotein M
    pept                 1 Similar to HHV8 Orf 40 - helicase - primase
    pept                 1 Similar to HHV8 Orf 41 - helicase - primase
    pept               (C1 Similar to HHV8 Orf 42
    pept               (C1 Similar to HHV8 Orf 43 - capsid protein
    pept                 1 Similar to HHV8 Orf 44 - helicase -primase
    pept               (C1 Similar to HHV8 Orf 45
```

FIG. 5

Comparison of Corresponding Repeats in RRV and KSHV

virus	insert name	total length	repeat unit length	G + C content
KSHV	*frnk*[1]	332 bp 292 bp	20 bp 30 bp	80.1% 84.9%
RRV	*syko*[1]	304 bp 1008 bp	26 bp 25 bp	53.3% 79.9%
KSHV	*zppa*[1]	308 bp 244 bp	23 bp 23 bp	74.0% 77.9%
RRV	*vrtgo*[1]	405 bp 1029 bp	19 bp 32 bp	74.6% 84.4%

virus	insert name	total length	repeat unit length	G + A content
KSHV	*mdsk*	409 bp	—[2]	75.4%
RRV	*brds*	196	13 bp	81.6%

1 KSHV *frnk* and *zppa* and RRV *syko* and *vrtgo* are tandem repeats.

2 KSHV *mdsk* is a complex repeat with no defined unit length.

FIG. 6

Comparison of Interferon regulatory elements coded by RRV and KSHVa

	KSHV K9	KSHV K10	KSHV K10.1	KSHV K10.5	KSHV K11	Rh R6	Rh R7	Rh R8	Rh R9	Rh R10	Rh R11	Rh R12	Rh R13
KSHV K9	100.000 100.000					26.044 21.130	28.291 20.728	28.857 19.427		33.705 26.184	29.972 21.849		
KSHV K10		100.000 100.000											
KSHV K10.1			100.000 100.000										
KSHV K10.5				100.000 100.000	32.036 21.895								
KSHV K11					100.000 100.000								
Rh R6						100.000 100.000		26.393 19.062	29.918 22.131	54.427 47.917			
Rh R7							100.000 100.000	34.513 26.254			50.773 41.495	33.038 24.484	
Rh R8								100.000 100.000		31.412 24.207	35.693 23.849	61.254 50.997	28.818 21.037
Rh R9									100.000 100.000	28.980 18.367	28.216 21.577	30.364 18.623	58.103 52.964
Rh R10										100.000 100.000	32.951 23.496	33.526 25.723	
Rh R11											100.000 100.000	33.923 23.849	31.124 25.072
Rh R12												100.000 100.000	
Rh R13													100.000 100.000

a Blank cells indicated no similarity; upper number is percent similarity; lower number is percent identity.

FIG. 7A

Comparison of RRV, KSHV and HVS ORFs

ORF	Strand	Size aa	KSHV Size aa	KSHV % Sim	KSHV % I	HVS Size aa	HVS % S	HVS % I	Putative Function
R11	+	423							
Orf 2	-	188	210	55.1%	46.0%	187	65.6%	54.8%	Dihydrofolate reductase
Orf 42	+	645	550	40.9%	35.7%	360	42.0%	35.3%	Complement binding protein
						287	44.0%	38.6%	
Orf 6	+	1132	1133	71.3%	63.3%	1128	65.2%	53.5%	ssDNA binding protein
Orf 7	+	686	695	60.1%	51.5%	679	58.1%	47.7%	Transport protein
Orf 8	+	829	845	73.3%	65.5%	808	62.4%	53.1%	Glycoprotein B
Orf 9	+	1014	1312	75.0%	67.0%	1009	71.0%	62.5%	DNA polymerase
Orf 10	+	384	418	43.5%	34.8%	407	33.6%	23.3%	
Orf 11	+	409	407	41.3%	31.7%	405	46.3%	32.4%	
R21	-	207							
Orf 70	-	333	337	72.1%	66.1%	294	72.1%	64.6%	Thymidylate synthase
R33	-	115	95	41.9%	32.3%				
Orf 16	+	187	175	58.0%	46.0%	160	31.4%	21.4%	Bcl-2 homolog
Orf 17	-	536	553	50.6%	44.3%	475	49.0%	42.2%	Capsid protein
Orf 18	+	299	257	68.1%	58.0%	256	60.2%	48.8%	
Orf 19	-	547	549	61.1%	52.8%	543	55.5%	46.9%	Tegument protein
Orf 20	-	350	320	51.8%	44.7%	303	43.2%	35.6%	
Orf 21	+	557	580	54.0%	44.6%	527	39.0%	31.7%	Thymidine kinase
Orf 22	+	704	730	50.1%	40.7%	717	42.3%	31.5%	Glycoprotein H
Orf 23	-	402	404	56.8%	48.5%	253	40.5%	29.8%	
Orf 24	-	732	752	66.3%	58.7%	731	56.3%	46.8%	
Orf 25	+	1378	1376	79.9%	72.5%	1371	76.7%	67.5%	Major capsid protein
Orf 26	+	307	305	71.8%	64.3%	304	69.1%	58.2%	Capsid protein
Orf 27	+	269	290	33.6%	25.3%	280	35.0%	27.1%	
Orf 28	+	91	102	30.1%	26.5%	93			
Orf 29b	-	348	351	77.6%	66.4%	387	74.4%	62.9%	Packaging protein
Orf 30	+	76	77	51.3%	38.2%	75	40.3%	29.2%	
Orf 31	+	217	224	56.0%	45.4%	208	50.5%	39.9%	

FIG. 7B

	RRV		KSHV			HVS			
ORF	Strand	Size aa	Size aa	% Sim	% I	Size aa	% S	% I	Putative Function
Orf 32	+	464	454	49.9%	41.8%	441	43.2%	34.1%	
Orf 33	+	336	312	52.1%	42.1%	330	49.1%	39.1%	
Orf 29a	-	327	312	66.7%	61.2%	303	57.8%	49.8%	Packaging protein
Orf 34	+	327	327	58.9%	48.5%	316	53.7%	40.6%	
Orf 35	+	149	151	47.7%	35.6%	150	51.0%	37.4%	
Orf 36	+	435	444	56.0%	46.1%	431	38.4%	28.7%	Kinase
Orf 37	+	480	486	72.4%	63.5%	483	63.0%	53.2%	Alkaline exonuclease
Orf 38	+	69	61	56.7%	45.0%	66	39.4%	34.8%	
Orf 39	-	378	399	73.0%	59.3%	366	67.1%	57.0%	Glycoprotein M
Orf 40	+	468	457	42.2%	32.7%	450	39.1%	28.1%	Helicase-primase
Orf 41	+	203	205	33.5%	26.0%	161	37.3%	29.1%	Helicase-primase
Orf 42	-	272	278	56.8%	46.1%	265	51.2%	38.1%	
Orf 43	-	576	605	69.7%	61.6%	563	66.4%	56.6%	Capsid protein
Orf 44	+	790	788	73.9%	66.0%	781	71.1%	62.6%	Helicase-primase
Orf 45	-	352	407	31.2%	24.9%	257			
Orf 46	-	255	255	71.9%	60.1%	252	67.5%	59.1%	Uracil DNA glucosidase
Orf 47	-	169	167	31.9%	27.7%	141	33.3%	23.9%	Glycoprotein L
Orf 48	-	389	402	36.2%	29.2%	797	34.1%	25.8%	
Orf 49	-	301	302	66.1%	54.2%	303	35.1%	23.3%	
Orf 50	+	514	631	46.6%	37.8%	535	29.7%	21.6%	Transactivator
R41	+	206							
R51	+	111							
Orf 52	-	139	131	58.5%	45.4%	115	41.7%	30.4%	
Orf 53	-	104	110	51.0%	46.2%	90	43.3%	28.9%	
Orf 54	+	290	318	48.6%	41.0%	287	46.5%	36.4%	dUTPase
Orf 55	-	210	227	62.9%	55.2%	200	52.5%	44.4%	
Orf 56	+	828	843	61.2%	52.5%	835	54.0%	43.6%	DNA replication protein
Orf 57	+	442	275	60.6%	47.1%	416	40.3%	31.5%	Immediate-early protein
R64	-	415		26.0%	21.1%				
R74	-	415		28.3%	20.7%				
R84	-	351		28.9%	19.4%				

FIG. 7C

	RRV		KSHV			HVS			
ORF	Strand	Size aa	Size aa	% Sim	% I	Size aa	% S	% I	Putative Function
R94	-	253							
R104	-	385		33.7%	26.2%				
R114	-	390		30.0%	21.8%				
R124	-	355							
R134	-	364							
Orf 58	-	360	357	45.2%	38.2%	357	39.9%	29.5%	
Orf 59	-	394	396	60.3%	51.8%	368	40.7%	32.7%	DNA replication protein
Orf 60	-	314	305	78.2%	70.0%	305	71.0%	62.4%	Ribonucleotide reductase, small
Orf 61	-	788	792	69.3%	61.7%	767	64.4%	53.3%	Ribonucleotide reductase, large
Orf 62	-	331	331	64.4%	56.5%	330	53.8%	41.9%	Assembly / DNA maturation
Orf 63	+	939	927	51.8%	42.6%	899	43.4%	34.6%	Tegument protein
Orf 64	+	2548	2635	49.6%	40.2%	2469	39.2%	29.4%	Tegument protein
Orf 65	-	169	170	48.2%	38.6%	139	41.0%	33.1%	Capsid protein
Orf 66	-	448	429	51.9%	46.4%	435	43.6%	32.3%	
Orf 67	-	224	271	69.6%	64.7%	253	58.6%	51.4%	Tegument protein
Orf 68	+	457	545	53.2%	44.8%	436	53.5%	44.3%	Glycoprotein
Orf 69	+	297	225	73.1%	65.5%	261	57.5%	49.0%	
R141	-	228							
Orf 71	-	174	139	38.8%	30.9%	167	25.3%	15.1%	Flip homolog
Orf 72	-	254	257	49.8%	38.6%	254	37.5%	29.2%	Cyclin D homolog
Orf 73	-	447	1162	23.6%	16.8%	407	29.0%	20.8%	Immediate-early gene
R155	+	253	348	35.2%	31.2%				
Orf 74	+	342	342	51.6%	41.1%	321	41.1%	32.1%	G-protein coupled receptor
Orf 75	-	1298	1296	52.2%	44.0%	1299	43.2%	34.4%	Tegument protein / FGARAT

% Sim., percent similar; % Id., percent identical; ssDNA, single-stranded DNA; FGARAT, N-formalglycinamide ribotide amidotransferase; 1, no similarity found; 2, compared to HVS ORF 4a and 4b; 3, compared to KSHV R4; 4, compared to KSHV K9; 5, compared to KSHV K14.

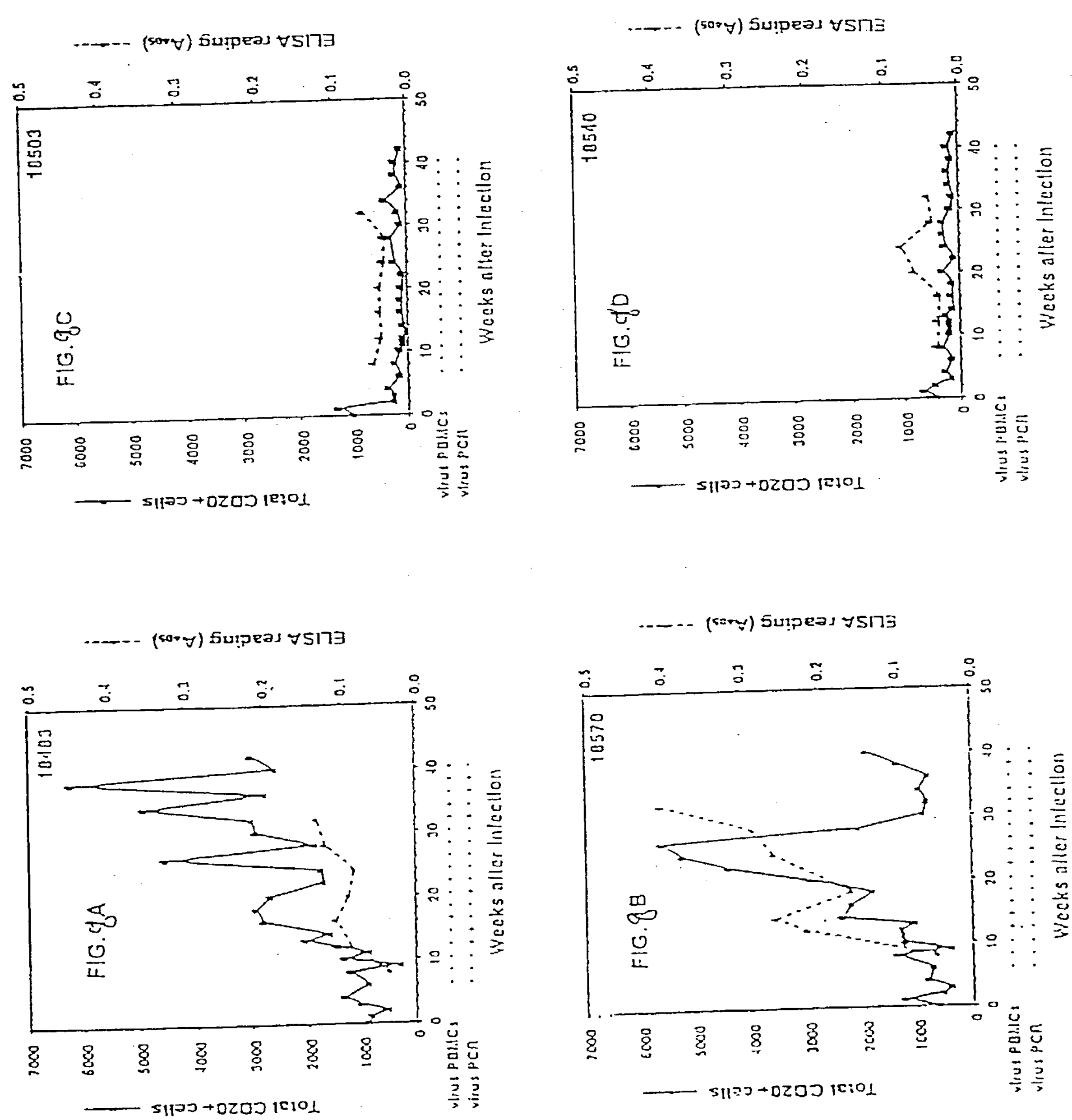


FIG. 8A
10403
FIG. 8B
10570
FIG. 8C
10503
FIG. 8D
10540
Total CD20+ cells
ELISA reading (A_{405})
Weeks after Infection
virus PBMCs
virus PCR

FIG. 8

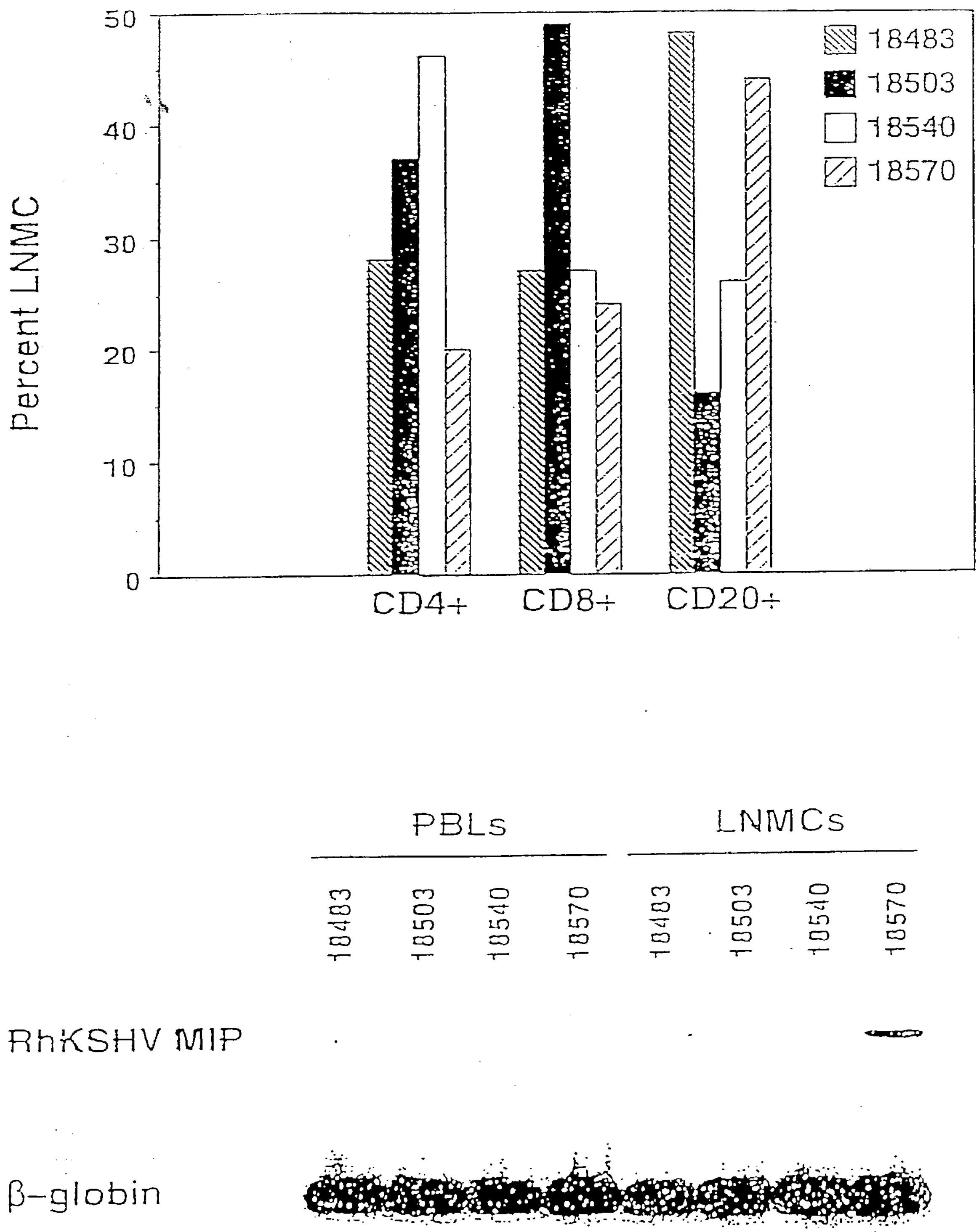

50
40
30
20
10
0
Percent LNMC
18483
18503
18540
18570
CD4+
CD8+
CD20+
PBLs
LNMCs
18483
18503
18540
18570
18483
18503
18540
18570
RhKSHV MIP
β-globin

FIG. 9

FIG. 10

```
atg ttc cct gtc tgg ttc gtc ttg ttt tac ctg tcg tgt tgg gcg gcc     48
Met Phe Pro Val Trp Phe Val Leu Phe Tyr Leu Ser Cys Trp Ala Ala
 1               5                  10                  15

agc cct acg ctg gcg cct ccc ccg act gcc gct gga att aac gtt ctc     96
Ser Pro Thr Leu Ala Pro Pro Pro Thr Ala Ala Gly Ile Asn Val Leu
            20                  25                  30

ccc cag tgg gcc ggc aac cgc gcc tct ctt gac agg acc agg ggg cgc    144
Pro Gln Trp Ala Gly Asn Arg Ala Ser Leu Asp Arg Thr Arg Gly Arg
        35                  40                  45

ctg tct gaa gtg ggg tta aac ata cag cgc tgg ttc gtt tac ctg tgc    192
Leu Ser Glu Val Gly Leu Asn Ile Gln Arg Trp Phe Val Tyr Leu Cys
    50                  55                  60

cac cac tcc act ctc tgt cgg gtg cgt gag tac ccg cgc atc atg tcg    240
His His Ser Thr Leu Cys Arg Val Arg Glu Tyr Pro Arg Ile Met Ser
 65                  70                  75                  80

ttt gtt cac ttc cct ata ttg atg tct aac gtt gag tgc cag cgc cgc    288
Phe Val His Phe Pro Ile Leu Met Ser Asn Val Glu Cys Gln Arg Arg
                85                  90                  95

gag ttt cgc ggg gcc gag tgt atg aac gcc atg gtt cgc ggg ctc cgg    336
Glu Phe Arg Gly Ala Glu Cys Met Asn Ala Met Val Arg Gly Leu Arg
            100                 105                 110

gcc tac gag agt tac ctg acg cga ctg agg atg ctg ctg gac gac gcg    384
Ala Tyr Glu Ser Tyr Leu Thr Arg Leu Arg Met Leu Leu Asp Asp Ala
        115                 120                 125

ccc ggg gac gcg gac gcc gcg gcc att ggc tcc gcg gtg acc gtg gtg    432
Pro Gly Asp Ala Asp Ala Ala Ala Ile Gly Ser Ala Val Thr Val Val
    130                 135                 140

ctg tcc gcc ctc gac tct cta att gag gag ctt ccc gta aat aac aag    480
Leu Ser Ala Leu Asp Ser Leu Ile Glu Glu Leu Pro Val Asn Asn Lys
145                 150                 155                 160

ata ggt ggc gcg gag tct aat gaa aaa acc gtg cgt gcg ttg gga ggg    528
Ile Gly Gly Ala Glu Ser Asn Glu Lys Thr Val Arg Ala Leu Gly Gly
                165                 170                 175

cag agc ccc cgg gac gtt gtt ctc agc gcg ttt cgc ata ctg gaa tat    576
Gln Ser Pro Arg Asp Val Val Leu Ser Ala Phe Arg Ile Leu Glu Tyr
            180                 185                 190

cta cag atg ttt ttg cgg gac ggg cgc cgc gca ata gct atg atg taa    624
Leu Gln Met Phe Leu Arg Asp Gly Arg Arg Ala Ile Ala Met Met
        195                 200                 205
```

FIG. 11

```
atg agg ggc ctt ttc gtg tgc gtt ttt ttt gcg gtg ttc gcg tgt gta    48
Met Arg Gly Leu Phe Val Cys Val Phe Phe Ala Val Phe Ala Cys Val
 1               5                  10                  15

gtt gat tat gcc ttt cct atg ggc tcc atg agc gga ccc gcg ccc gaa    96
Val Asp Tyr Ala Phe Pro Met Gly Ser Met Ser Gly Pro Ala Pro Glu
            20                  25                  30

ctc tgc tgt ttg ggg tat gta act cat ctg ccg cca ccc ggt tta gtg   144
Leu Cys Cys Leu Gly Tyr Val Thr His Leu Pro Pro Pro Gly Leu Val
        35                  40                  45

gtc tct tac tcc cac acc tcg tcg cag tgc tcg gtg gac gcc gtg ata   192
Val Ser Tyr Ser His Thr Ser Ser Gln Cys Ser Val Asp Ala Val Ile
    50                  55                  60

tta aac act cgc cgc ggt aaa aag ctg tgt gcc aat ccc ggg gac gac   240
Leu Asn Thr Arg Arg Gly Lys Lys Leu Cys Ala Asn Pro Gly Asp Asp
65                  70                  75                  80

gca gtg aag aaa ctg ctt cag gcg gtg gac aag cgt ccc aaa aag ggc   288
Ala Val Lys Lys Leu Leu Gln Ala Val Asp Lys Arg Pro Lys Lys Gly
                85                  90                  95

aga aga acc cgg cgc agc ctg att gac gat tcc gaa gag ggc ctt ggc   336
Arg Arg Thr Arg Arg Ser Leu Ile Asp Asp Ser Glu Glu Gly Leu Gly
            100                 105                 110

agc ggg att tag                                                   348
Ser Gly Ile
        115
```

WDN/SAS:gtc 4/26/01 178-59010

COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled CLONING OF RHESUS MACAQUE RHADINOVIRUS GENOME AND METHODS FOR ITS USE, the specification of which

☐ is attached hereto.

☐ was filed on _____ as United States Application No. _____.

☒ was filed on November 5, 1999 as International Application No. PCT/US99/26260.

☐ and was amended on _____ (if applicable).

☐ with amendments through _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56. If this is a continuation-in-part application filed under the conditions specified in 35 U.S.C. § 120 which discloses and claims subject matter in addition to that disclosed in the prior copending application, I further acknowledge the duty to disclose material information as defined in 37 C.F.R. § 1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT International application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT International application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

Prior Foreign Application(s)			Priority Claimed	
			☐	☐
(Number)	(Country)	(Day/Month/Year Filed)	Yes	No

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below:

Application Number	Filing Date
60/107,507	November 6, 1998
60/109,409	November 20, 1998

WDN/SAS:gte 4/26/01 178-59010

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or § 365(c) of any PCT International application(s) designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, § 1.56(a) which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

PCT/US99/26260	5 November 1999	Pending
(Application No.)	(Filing Date)	(Status: patented, pending, abandoned)

The undersigned hereby authorizes the U.S. attorney or agent named herein to accept and follow instructions from _____ as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney or agent and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney or agent named herein will be so notified by the undersigned.

I hereby appoint the practitioners associated with the customer number provided below to prosecute this application, to file a corresponding international application, and to transact all business in the Patent and Trademark Office connected therewith:

Customer Number 24197 KSCLW

Name	Reg. No.	Name	Reg. No.
BLYVEIS, Deborah B.	47,337	PETERSEN, David P.	28,106
CALDWELL, Lisa M.	41,653	POLLEY, Richard J.	28,107
GIRARD, Michael P.	38,467	RINEHART, Kyle B.	47,027
HAENDLER, Jeffrey B.	43,652	RUPERT, Wayne W.	34,420
HARDING, Tanya M.	42,630	RYBAK, Sheree L.	47,913
JAKUBEK, Joseph T.	34,190	SCOTTI, Robert F.	39,830
JONCUS, Stephen J.	44,809	SIEGEL, Susan Alpert	43,121
JONES, Michael D.	41,879	SLATER, Stacey C.	36,011
KLARQUIST, Kenneth S.	16,445	STEPHENS Jr., Donald L.	34,022
KLITZKE II, Ramon A.	30,188	STUART, John W.	24,540
LEIGH, James S.	20,434	VANDENBERG, John D.	31,312
MAURER, Gregory L.	43,781	WHINSTON, Arthur L.	19,155
NOONAN, William D.	30,878	WIGHT, Stephen A.	37,759
ORR, David E.	44,988	WINN, Garth A.	33,220

JL

Address all telephone calls to William D. Noonan, M.D. at telephone number (503) 226-7391.

Address all correspondence to:

Customer Number 24197 KSCLW

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

WDN/SAS:gte 4/26/01 178-59010

1 - 00 , Full Name of Sole or first Inventor: Scott W. Wong

Inventor's Signature [signature] 4/27/01 Date

Residence: Beaverton, OR

Citizenship: United States of America

Post Office Address: 7926 S.W. 167th Place, Beaverton, OR 97007

2 - 00 Full Name of Second Joint Inventor, if any: Michael K. Axthelm

Inventor's Signature [signature] 4/27/01 Date

Residence: Hillsboro, OR

Citizenship: United States of America

Post Office Address: 1345 S.E. 35th Avenue, Hillsboro, OR 97123

3 - 00 Full Name of Third Joint Inventor, if any: Robert P. Searles

Inventor's Signature Robert P. Searles 4/31/01 Date

Residence: Beaverton, OR

Citizenship: United States of America

Post Office Address: 18300 N.W. Heritage Parkway #106, Beaverton, OR 97006

SEQUENCE LISTING

<110> Wong et al.

<120> CLONING OF RHADINOVIRUS GENOME AND METHODS FOR ITS USE

<130> 53683

<140> 60/107,507
<141> 1998-11-06

<140> 60/109,409
<141> 1998-11-20

<160> 179

<170>

<210> 1
<211> 133719
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

```
   1 GATCGGGAAA ACGCGAGGGG AGCGGGGGAC AGGGGACGGC GTGTGCGTGC TTGTGAGACA
  61 CCGGGTACGG CTGCCTGCCT GCTCGCTGGC CTGCTTGCTG AGGGGACAGT AGGCCTGCTT
 121 GCTCGCTGGC CTGCTTGCTG AGGGGACAGT AGGCCTGCTT GCTGAGGGGA CAGTAGGCCT
 181 GCTTGCTCGC TGGCCTGCTT GCTGAGGGGA CAGTAGGGCT GCTGGCTTGC TAGTAGGGCT
 241 GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTT
 301 GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTT
 361 GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTT GCTCGCTGGC CTGCTTGCTG
 421 AGGGGACAGT AGGGCTGCTT GCTTGCTAAG GGGACGGTAC GCCTGCCTGA TGGCTTGATA
 481 GTAGGGCTGC TGGGCTGCTA GTAGGGCTGC TGGGCTGCTA GTAGGGCTGC TGGGCTGCTA
 541 GTAGGGCTGC TGGGCTGCTA GTAGGGCTCC TGGGCTGCTA GTAGGGCTGC TGGGCTGCTA
 601 GTAGGGCTCC TGGGCTGCTA GTAGGGCTGC TGGGCTGCTA GTAGGGCTGC TGGGCTGCTA
 661 GTAGGGCTGC TGGGCTGCTA GTAGGGCTGC CTGCTGGCTT GCTTGCTTGC TTGCTAGTGG
 721 GGCCGCTTGC CTGCTACTAG GGCTGCTGTG CAGCTGGGAG AACAGAGTAG GGCTGCCGGC
 781 CAGCTGCGTG CGAGGGCGTC CGAGGGCCAG ACGAGGACAC GGGACCCGGG CCTCTCCCCC
 841 AGGCACAAAG CAGTAGGGCT GGCCAGGGGA AACAGTAAGG CTGATTGCTT GCTGAAAAAC
 901 AGTAGGGCTG CTGGTTTGTT GCTAACGGAA AAAGGGGAGG TGTGTGTATG CTTTTGTGAC
 961 TTTCCAGCGG AGGGGGTACA GTGCACGGCC AAGTTACAAG CACCTGCTTA ACTTGCTTTG
1021 GCTCTGTGCG GTTTTGTTGC TAGGTACTAG TAGTAACACA TAGTATTTCA TCGCATGGCG
1081 CCTGCATACG CTCCACAGCA TCCGAAACAC GTTATTTCTA TAGACTAACT TTAGTGTTCT
1141 TCTTGGTTGG TACCATTTTA CGAAGTTTGC TCCGTTGAAT AATAAGAAAT CCGTTGTGGT
1201 TACAATACAC CTGCCAACGT ATTGGATGGT TCCTTTTGCC ATCAACCTTT GCTTGCATTC
1261 TAACATTGAA TTTTCACATT ACACCCTAGT CTTCTCATGC AATTAGGTAT GTCTTTGCTC
1321 GTGTTTAATT TTCTCTACAC ATTTCTGTGG AAATGTTTGT GTTGGTTTTA TTTATGTTAT
1381 TGCAACCTGT ATCGGTGGAG CTTTTGCCTG CAAAATTAAC TTCTGTTCCA ACGTGGTGTC
1441 CACCACATCC CGGAGATACC TACTTGCTAA CCTGCCGCGG GACGTCTACG GCCAGAGACC
1501 AGCGAAGCAC ACAATGGTTT CGCAACAACA CGCTTATGCG TGGGAGTAAT TTCTACGGCA
1561 GACTGGTATC TGTGACTCCC AATGCTACGA TATCTGACCG GTATGCGTGT CAAACAAAAA
1621 CAACAACGCG GAGTAACAAC ATCGATTTTC GGGTAAGCTC ATCGCGCCTC ACGCTCCAAG
1681 AACGGTGCTC TTCATACGGC TATACTTACG CGAATAACAC AAGGGTATTG AGGTGTTACT
1741 CTGGTGGAAA CGTAACTTTA AGAAACGTTG TCTTTCATTT AAACGGTACA GCGGTCATCA
1801 ACGGTACTAC AACAAACATA CATACATTTG TGTTAACAGA AAAGACAGGA GGGACGTATT
1861 TCTGTTCTGC GTTTATTGGG AATGAAAAAT TCTATTCTCA GACAATTAAT GTGTTTTTTA
1921 CTTCATTTAC CTTTAAACCT ACAAACGACA TTCCCAATGA GTCACATTTT AATAAAACTG
1981 GGCAAATACA ACAAACAGCT AGTGTACAAC ATCCTGAAAA CTACGTTGTG TTCTCTGTTC
```

2041 CCGTTTTTTC TATTGGCGTT TTAACAGGTA TTGCAATATC GTTGATTATG TGTTGGTTAT
2101 TTACAATACG CTGCAACGAG AACTCTGAAT CATCAACTAA TAGTTATGCA AGCCAGACAA
2161 GCTACATTCA ACCCTCCCAT AATCAGCGTT CCAATACTAA TGAATGTAGT CGCCATACCT
2221 ACAGAAATGC TCATCAAGAA GAGAGTATTG AAGAACTACC AAACCAACAC ACAAGTGAAA
2281 CTGATTCTTG CTGTCAATTA GTTTTACTTG AAGTGAAAAA TGTAGCCTAC GATGGACCGC
2341 AGGAAAACAC AATTAACGAA GTTATGGAAC AGTATGATGA TGTGGTTGTA AAAAATATAG
2401 AACAAACATC ATATGAGGAT AATGTTGAGC ACATGGACTA TAGTGATACT ATAAATCCCA
2461 ATTTTAATTA CTACAGTGGA CTAATATTGG AAGAAGTAGA TGAAGTTTTT TACAATGAAC
2521 TAGAAAATCA ATATCATGGA TTAATACTGG AGAATTTAGA TCACAATGAG TACAATCATT
2581 TAAATGAATT AAACATGATA GAACAATATG ATTGGTTAGA ATAAATAACT TGTGGGTTAT
2641 TTTTAAATTA AAACATAAAC AACATTAAAG CAGCATTTTT GTGTAAATCC TTTATTTATT
2701 AAAATTTTTT TCATATACCT GAAACTTATA TTTAATTCCA TTTTCCTCAA AGTTTGTATC
2761 TTTGCCCGGT ATCTCTGATA ACATGGTGTA TTCGGTAAAA TTAATTGATG GAAAAAATAC
2821 ATCACAATCA AAACTCTCCA TAATGCGCGT AATATATAGT TTTAAGGGAC ACTTATAGTT
2881 AAGAACACTC TCATAAACAG ATTTTCCTCC AATAACCCAG ACCGTGTTTA ATTGTTCTTT
2941 TAGTTTGTAT TGTCTATAGA AATTAAACGC ATCGTCTAGC GTTCTCGCTA GAAAGTGTGC
3001 TCCGTGCGGT GGTTCACGTA ATTCTCTGCT CAAAATAATA TTAATTCTGT TCACCAGAGG
3061 GCGCTTCTTT TCAGGGATGG AAAACCATGT TCTTTTCCCC ATAATAACCA CATTCTTTTC
3121 ACCTACAACC GATGGCGTAG ACGTCATTTT CTGAAAATAC ATCATTTCGT TTCTAAGGTA
3181 CGGCCAAGGC ATTGTTCCGT TTTTACCAAT TCCTAATTGT TCATCAACTG CAACGATGCA
3241 GTTAACTGTA ATGTCCATGG TTTTAGTTGC CACTGACGGG TTTAACACAG AAGTATTTCA
3301 CAATTATAAA CAAATAACCC ACGTGACATG TACTTACTAA TGTAAGTACG TAACGTGATA
3361 TAGTTAATCA TATTCGCTTT CACATAATGG ACAAAACGAA AAAATGTTAC CTGTGTTAAG
3421 CCTACCTTGT TTGGATGGAT TTATGGTTAC AATGAAAAAT AAAAATATAT ATATAACGTT
3481 TTTTACGTAA AACATATTGT AATTTAATCA CGTGATGTTT ATGTTGGCTA CTAACACACA
3541 ATGTTTGCAT AAAAGCACTT ATGTTAGTGA GCTTCAATCT TTATATTTTA TTCAAATTGT
3601 TTAAAGAGGC ATGTGAATAT ATTTAACACG CATTCAGAAA TAGGCGCAAC GCTGTGCCGT
3661 ATAACTGGTA AAAACATGAC GTTTAAACTT TTTCCTCTGT TTTTATTACA CGCCATAATG
3721 TACGTCCACT GCGATGAAAA CTGTAAACCT CCACATTTCA CGGAATATCG CGTCAAGTCT
3781 AACACAGAAA AGGACTTATA TAGTGTTGGA GAAACAGCTG AATTAATTTG TCGTCCTGGT
3841 TATGTTACAA ATACAAAAAT AATAACAACA GAATGTTTAC AAAATGGTAC GTGGTCAACA
3901 CCAAATTTTC CATGCGACAG AAAAAGATGT CCCACACCTG CTGACTTGCT GAATGGAGCC
3961 GTGCACATTC ACGGGGGAGA TAATGCCTTA AAATTTGGAT CCAATATTTC CTATGAGTGT
4021 AATGAAGGTT ATGATTTAAT TGGTAGTAAT GTTCGTTTTT GTATTTTACA AGACACAGAA
4081 AACGTAAATT GGGATTCAAA TGAACCAGTC TGTGAAATTC AGAAATGTAT TAAACCACCG
4141 GCAGTGGAAC ATGGGGACTA CCTACCTAAC CAAGATGTTT ATAACTATGG AGATGCAATT
4201 ACATTTAAAT GTTCATTGTC GTATACACTC GTTGGATCAA CAACATTAGT ATGCACGTCA
4261 AACAAAAAGT GGTCAAACTC TTTCCCAACG TGTTTAATGC TCGTATGTGA AAGTCCACAA
4321 ATAGACAATG GGTACATAGA CATTGGCTTG TCCAGAAGAT ACAACCATGG ACAATCAATT
4381 ACTGTAAAGT GTAGCGACGG GTACAACATT GTTGGGCCTG AAACATTAAC GTGCACAAAC
4441 ACAACTTGGG TTCCACCATT ACCTAAGTGT GTATTAGTTA CAAATAACCC AAGCACACCC
4501 ATGCCAGAAA CACCCATGCC AGAAACACCC ACGCCAGATT ATCAAAAAAT AAATTTGTCA
4561 ACCGCTAAAA CTGCAACAAC ACCAAATGCG TTTGTTACAA CTGTTGTTTC TCCAGAGAAA
4621 GACGACGTTA CTTGTGTAAA GCCTCATTTT GAGCGATTCA TGGTAAAGGC TGAAAATGAC
4681 AAGGAAAAAT ACAGTGTTGG TGCGAGCGTT GAGCTAATAT GTCGACCAGG ATTTACTAAA
4741 ATGCAGTCTA CAGTTTCTGT TGAATGTTTG TCCAACGGAA CATGGACTGC TCCAAATGCC
4801 AAGTGTCATA GAAAAAAATG TCCAACCCCT CAAGAACTTT TAAACGGAGA GTATATAGTT
4861 ACAAGCGGAG AAGATGCTTT TAAGTACGGA ACAAATATAA CATATAAATG TAATGAAGGT
4921 TATCAACTTT TAGGAAGTAT GGTGCGGATT TGTATGCTTA AAGACGATTT AAAAACAGTT
4981 GACTGGGAGC CAAAAGCGCC TATATGTGAT ATTGAAAAAT GTAAGCCACC GCCACAAATT
5041 ACAAACGGAA AATACCATCC GGTGAAAGAC TTTTATCAGT ATTTGGACAC CGTAACATTT
5101 TCGTGCAATC GTGACTTTTC TTTAGTTGGA GATGAAATGA CAACGTGTAT AAGTAATACG
5161 TGGAATAAAC CGTTTCCAAG ATGTGAACAA ATCACTTGCA GCGCTCCTAA TATTGCACAC
5221 GGAAAGCTGC TAACAGGTTC TTCAAGCGTT TACAAATACG GTCAATCTGT TACCATTGGT
5281 TGTGAAACTG GATTTACTCT AATTGGCAGT GAAATTTCTA CATGCAAGGA TTCATCGTGG
5341 GATCCACCAC TTCCTACGTG CGTGCCAGCT GTTTCAATGC CTTCTGACAC ACCTAAACCA
5401 GAAACCAAAA AACCAAACAC GCCAACGCCA GAAGCACCCA AACCAAACAC CCCAAACGTT

5461 GGAACACATA CACCATTCAA ACCACCACCA CAAAATCCAC CAATAGCACC CCCAATGAGT
5521 AAATGGAAAA GGCATGTCGT GTTAGTTCTT TTTGCAAGTG TCGCGTCCTT GTTATTCGTA
5581 CTTGCTGCCC TTTATTGTTG TTTTCTAAAA TAACTGTTTT TTGTCTTCAG CAGGTTCGCC
5641 AGGCAAACTC GCACGCATTA ACCAATCTGC CAACCGCCGT TGATTCCGGA ATTAAGTTTA
5701 CATTATTCAA GGTTGCCAAT AAAGGTGGTT TAAAAATATT CTATTGGTGT TCATTGTTTT
5761 ATGTTGACCC GTTTATAGTT ATCGCGCCAC CTTGTGGCTA CATTATATAG CACGATCACT
5821 TTCCACGTTA TACTTTCACG TACTATGACT CATACGCCTT AACGTCACGT GGCGTGCGAT
5881 TGTGGCCGGG GCTGAAAATA ACACAAGGGG TACATAATCC ATCCAGGCGG CACACATTAG
5941 ACACGGTTTA TAAAACTATA TCGGATGCGC CAACAATCAC TGTCGCTAGC GACACTGATA
6001 GAAAAACATT TTAACGTTTG TTTAGCGAAC TTGAATAACA CATAATGGCT TCCAAAGGCA
6061 ACGCCGGACA ACCCCTGGAA GATAATCAGG GGTCTCGTGC CCCGATAGGT GCGTGCGGAT
6121 ACGTGTACGC GTATTCGAAA CAAGACTTTC CCTTTGCCGA GGCGTCCATA CTCGGCAACA
6181 GACCATCTGG ATCTGGCGTT TTCTCGCTAC CAATCCTTTA CGGACTTACA GTTGAACACG
6241 AATTCCCTCT CACCGTAAAA GCCGCATACA AAAAAGTTGA CACCACGACG CTCGCCGTTA
6301 AGGTGACGTG CTTTCACAGA GAGGTTATTG TGTTTCACAA TGCAAGTTTA TTCAGGCCGG
6361 TGTTTGACGG TACCGGTCTT AACGAACTAT GCGAGGAAGC CAGGGCTCTC TTTGGGTACA
6421 CGCAGTTTAT AGAACCGGGT CCACCTCACA GCATATGGAA CCCTCTGGAA TGTCCGCAGT
6481 TACCGGACAA GGATGAGATG TTTCTTGGCG TTGTTGTTAC GGAAGGGTTT AAGGAAAGAC
6541 TGTGGAGGGG CTGTCTCGTT CCCGCGGTGT TCCAGACCCA GCAGGTGCAG ATTGCCGGAC
6601 GCCAGGCGTT TAAAGTGCCG TTGTACGACG AAGACCTGTT TGCACCTCAC GGTCATAGAA
6661 TGCCAAGGTT TTACCATAAA GACGTTAGCG CGTACCTCTA CGACTCCCTC TTTACCAGCA
6721 TCGCCCAGGC CCTGAGACTC AAAGACGTGA CGGCGGTCAT CCACGCCACA GAAAAGCAAT
6781 TCATGCAGGA CCATTACAAA ATTGCCAAGA TAGTGCAGGC AAAACAGTTT TCAACGACGC
6841 TGCCGAAAAC GACAGACGGG TCGTCCCACA TGATTGTGGA CAGCGTCGTC GCCGAGCTCG
6901 CCCTTAGTTA CGGCTGTATG TTTCTCGAGT GTCCCCAGGA CGCGTGCGAG TTGCTGAACT
6961 ACGATAGCTG GCCCATATTT GATGGTTGTG ACTCACCAGA GGCTAGGGTT AACGCGTTAG
7021 AGCGCTGGTC GGCCGAACAG GCCGTTCACG TGGCGGGTCA ACTGTTCGCT GCCAATTCGG
7081 TGCTGTACCT AACTAAAGTG CAGAAGCAAG CGCCCAGGGG ACAAAAGGGA GACGTAAACG
7141 TGTACAACTC CTTTTTCCTC CAACACGGAC TGGGGTTTTT AAATGAGGCC ACGATCAAGG
7201 AAAACGGCAG CGAAGCCTTT AAGGGCGTAC CCTCAAACGC CCTCGATGGT TCTTCGTTCA
7261 CGCCGTATCA CCTGGCCTAC GCCGCGTCTT TCTCGCCCCA TCTGCTGGCG AAGTTATGTT
7321 ATTACATGCA GTTCTTGCAA CACCACAAAA GCTCCACGAA CCAGGCGTTT AACATGGTCC
7381 ATTATGTCGG CACCGCCGCC AACTCAGAGA TGTGCACGCT ATGTCACGGC AACACGCCGG
7441 CAACGTGCCT CAACACGCTG TTCTATAGAC TGAAGGATAG GTTTCCCGCC GTAACCACCC
7501 CTCAGCGCAG GGACCCCTAC GTGGTGACCG GAACAGCCGG GACCTTTAAC GACCTGGAGA
7561 TTCTGGGCAA CTTCGCGAGC TTTAGAGACC GCGAAGAGGA CGGAAACCCG GCCGACGAGC
7621 ACCCAAAGTA CACGTACTGG CAGCTATGTC AGACCGTGAC AGAAAAGCTA TCCGCGATTG
7681 GAATCACCGA AGACCACGAT AATCACGTGA ACCTCATCAC CAACATCCAA AGTTTTCTCA
7741 GGGTGTTCAA GGGTATCGAC TCAATTGTGG ACGGAGAGGT CATGAAGTTC GTTAATTCGA
7801 TGATTAAAAA TAACTTTAAT TTCCGCGAGC ACGTCAAATC GGTCCATCAC ATACTCCAGT
7861 TCTGCTGCAA CGTGTATTGG CAGGCGCCGT GCGCGGTGTT CCTGAATCTG TACTACAAAT
7921 CCCTGCTGTG GATCATTCAG GATATCTGCC TGCCGTACTG CATGATCTAC GAACAAGATA
7981 ATCCGGCGAT GGGCATCCTC CCCTCCGAGT GGCTAAAGAT GCATTTTCAG ACGTTGTGGA
8041 CGAACTTTAA AGCGGCGTGT CTCGACCGCG GGGTCCTCAC GGGGTGCGAA CTGAAAATCG
8101 TACACCGGGA CATGTTCTGC GACTTCTTCG ACACCGACGC GGGGTCCAAC GGCTTAATGG
8161 CCCCCTTTAA AATGCAGGTT AGAATAGCCC GAGCCATGAT GGTCGTTCCG AAATCAATTA
8221 AAATAAAAAA TAGAATCATT TTTTCCAACA CCGCGGGATC CGAGGCGGTG CAGTCGGGGT
8281 TCGTCAAACC GACGGGAACC AGGGACACTT ACGTGGTGGC CGGACCGTAC ATGAAGTTTC
8341 TCAACTCGCT GCATCGCGCG CTGTTCCCCG ACACCAAGAC CGCCGCGCTG TACCTGTGGC
8401 ACAAGATCTC CCAGACCAAC AAAACCCCAG TTCTGAAAGA CGTCCCGGAC GACGAGCTGG
8461 CGGAGCTGGT GTCGTACGTA AAGACCAACA GCCTCGCGTT CGAGGAAACG AACGTGCTGG
8521 ACGTGGTTCC GGATTCACTC ATGTCGTACG CGAGGATCAA ACTGAACGGG GCCATTCTAA
8581 GGGCATGTGG CCAGATTCAG TTCTACGCCA CGACGCTGCA CTGCCTCACG CCGGTGCTAC
8641 AGACGATCGA TGCCGAGGAA TACCCCCACG TGCTGGGCTC CGCGGCAATC GCCACACCGG
8701 TGGCTTACCT GGCAGAAATA CGCGGCCGCA CCGCCCTCAC CGTCCAGACG ACGGCGCGTC
8761 AGCCGGTCGC CGCCACAGGG CGCCTGCGTC CCGTGATAAC CGTTCCCATG GTAGTCAACA
8821 AATACACGGG GGTCAACGGG AACAACAACG TTTTCCACTG CGGAAACCTG GGGTACTTCG

```
 8881 CGGGGCGCGG CGTGGACCGC AACCTGTGGC CGGAAAGCTC CCCCTTTAAG AAAACGGGCG
 8941 TCAGCGCCAT GCTAAGAAAG AGACACGTCA TGATGACCCC CATTATCGAC CGCCTAATAA
 9001 AGCGAGCCGC GGGACAGACA ATCAGCACGT TCGAGGCGGA AAGCGTTAAA AGGAGCGTGC
 9061 AGGCGCTGTT AGAGGATAAG GACAACCCTA ACCTATTGAA GTCGGTAATC TTGGAGCTTA
 9121 TACGACACCT GGGGAAGGGC TGCCAGGACT TAAGCTCCGA GGACGTGCAA TATTACCTCG
 9181 GTGACTATTG TATGTTGACG GACGAGGTTT TATTTACGTT GGATAATATA GCACAGTCAG
 9241 GCGTGCCGTG GACTATCGAG GACGCGGGTG CCCTAATAGA GGATCGCCAG GACGCAGACG
 9301 ATCTTCAGTT CGTAGACAGC GACGATATCG CCACCGCTTC CTGTCAGCCC CCCGAGGAAC
 9361 AGCTACCGAC CCCTAGCGCC GGCGCCCTAC TGGCCGGGAA GAAGCGAAAA ATTAACGCGC
 9421 TGCTGAGCGA TCTAGACCTT TAGGAAAACC GTGGCAGGCG GGCAACAATG GCCAGGGAAC
 9481 TCGCAGCATT ATACGCGCAG CTGTCGGCCC TCGCCGTCGA CTTGAGTCTG GTTATCTTTG
 9541 CGGACCCGCG AAGTATCGAC GGTGCCCGCA TTCTAAAAAC AAAAACACAG ATAGAGAACC
 9601 TGAACCGCGA CCTTCTGCCG CTGCTACGCG AGCAAAACTC GGTAGAGACG TCCAGCCTGT
 9661 CGCTCGAAGT GGAGCACCTG GCCAAAAACA TCGAGGACAA ACTCGGCGAG CTGGAGCGCA
 9721 GTCTGCGGCA GAGATATTCG AGCCGAGAGC ATTTTGAAAC ACTACACCTG AGACCCGAAT
 9781 GTCACTATCA CTCTACGGTT ACTTTTCAGT TTTACGGGGG CGGGTTAATA GATGTAAACA
 9841 TGTGCCTAAT AAACGATGTA GAACTGCTGT GTAAAAGACT AGGGAGTGTG TTTTATTGCA
 9901 TCGGTGCAAA CGAAGCTCTG TCCGGATTGA ACCGGGTTCT GACGTTTCTG TCAACACTGC
 9961 GGGGTATCTC CCCGATCCCG CACCCAGACC TATACGTCAC GTCAGTGCCT TGCGTACAGT
10021 GCCTGAGGGA AATCGAACTC GTACCAAATC AGGGGTCCAG TTTACTCGCG GTGTTGGCAG
10081 ACCGACACTG CGATCACCTC TGTAAGAAGG TTAGGGCGGA GCCAATACAC GGCCTGTTTG
10141 AGACAGAACT GAGCCAGCTG GGTCTAAAAG TAACAAAACG TTCGGACGCC ACGCAGCACG
10201 GCGTCCGGTC CTCTGCAGAT CAGTTAAGGG AGTCGTCGCT GGCGGCCATA CAAGATCACA
10261 ATATATTCAA ACGGGTGTCC GCGTCAATCA TGGAACTATC CAATCTAATT TATTGGAACG
10321 CCGGGCAAAC CGGCCTCCAG ACCGGGACCG AAAACGAGTG CTCACAAATG GCCAGACTGC
10381 TAACACACGA GGCCGATATG CACGAGCACC GTGCGCTAAT AACACCCAAA CTAAGCGCGA
10441 CTCACTTCTA CGACTGTTTC CGACCGGATC CCATAGAATC CCTGTTCTGC GGCGGTCTTT
10501 TTAACTCTAT AGACGACACC ATAAACGCAC TGAGCCGGGA TTGCTCCGTG ACGTTCTTTC
10561 AACAGGCAAA CTATACCAAC GTTATGCGAA AACAAAACGA GCTGTTCACC AGACTCAATA
10621 GCATCCTGCG TCAGGGGAGC GCGGGATCGC AAAAACCGGC CACCCCCTCG GAGCCACGGA
10681 CCACCACCGT GGCGGCAACC GCGGCAAGCG ACGTCATCAA AGACGCACAG TATCGCAAAG
10741 AACAGTACAT GAAAAAGGTG GCCAGGGACG GCTTTAAAAA ACTAACAGAG TGTCTGCAGA
10801 CGCAGAGCGC GGTGTTGGCA AACGCACTCT GCATGCGCGT ATGGGGGGGC GTCGCCTACG
10861 GCGAGGCGTC CGAGCTGGTG AACCATTTTC TCCTCAGGCG GCGCTTCGTC GCGCTTCCCT
10921 GGGAGGCGCG CTGCCGCTCG GATCAGATTT TATTCGAAAA CTCAAAGTAC ATTAAAAACT
10981 CACTATATTC CCAGCGCCTC AGTCGCGAAC ACGTAGAGAT TATCACGCTG CAGTTTTACG
11041 GCCTGATAAC CGGCCCCCTG ACGCGCCAGA GCGATCTCTT TCCCGGCCCC GCCAACGTCG
11101 CGCTGGCCCA GTGTTTCGAG GCGGCCGGAA TGCTTCCGCA TCACAAGATG CTGGTGTCAG
11161 AGATGATATG GCCCCAGATT CAACCGAAAG ACTGGATAGA CCAGACATTT AATCGTTTTT
11221 ACCAACTTCC CGAAGGTGAT CTCAACGCGG TACAAAAGTC CGCCTGGTGC TTTATACGAG
11281 AGCTCGTCCT CTCGGTGGCG CTTTATAATC GCACGTGGGA AAAGACGCTG CGGATATTTT
11341 CCCTAGCGCG CGAGAAACTC TCCATCTCCA ACCTAGACGT TAAAGGCCTG ACGTCCGGCC
11401 TGTATCTAAC GTACGAGCAA GACGCGCCGC TCGTTCTAAT TTCTCAAAAT ACCGGCTGGA
11461 TATTTAAAGA CCTGTACGCT CTTCTGTACC ATCACCTGCA ACTGTCCGAC GGCCATGATG
11521 ATAACTAACC GAACGCGTCG TCTCCTGCGG GCGTGGGTCG TGATAATCGC GATCGGCACG
11581 GCGGTTGGCG AAAACGTCAC CACCCCCAAG GGCGCGACCA CCACCGCGAA GCCAACGCCG
11641 GGCCCGTCGA CGCCCACACC TCCCGAGAAC CCACCTAGGG CCGAGGCGTT TAAGTTTCGC
11701 GTGTGCAGCG CCTCGGCCAC CGGCGAACTC TTCAGGTTTA ACCTGGAAAA AACGTGTCCG
11761 GGCACCGAGG ACAAGACGCA CCAAGAAGGC ATCCTGATGG TGTTTAAAAA AAATATTGTC
11821 CCGCACATCT TTAAGGTCAG ACGGTACCGC AAGGTGGCCA CCTCGGTGAC CGTCTATCGA
11881 GGGTGGACCG AGACCGCCGT GACCGGCAAG CAAGAGGTCA TCCGACCGGT GCCGCAGTAC
11941 GAGATCAACC ACATGGACAC GACCTACCAG TGTTTCAGCT CCATGCGCGT AAACGTCAAC
12001 GGCATAGTAA ACACCTACAC GGACAGGGAC TTCACTAACC AGACCGTGTT TCTGCAACCG
12061 GTCGAGGGGC TCACGGATAA CATCCAGCGA TACTTCAGTC AGCCGGTGCT GTACACGACA
12121 CCGGGATGGT TTCCGGGAAT TTACAGGGTC AGAACCACGG TCAACTGCGA GATCGTGGAC
12181 ATGATCGCGC GTTCGGCGGA ACCGTACTCG TATTTTGTCA CCGCCCTGGG AGACACGGTA
12241 GAGGTGTCCC CGTTCTGCCA CAACGACTCA ACGTGCTCGG TCGCGGAGAA AACCGAAAAC
```

12301 GGCCTCGGCG CCCGCGTGCT CACAAATTAC ACCATCGTCG ACTTCGCGAC CCGCCAGCCC
12361 ACCACCGAAA CGCGGGTCTT CGCCGACTCG GGAGAATACA CCGTATCGTG GAAGGCGGAG
12421 GACCCCAAGT CGGCGGTCTG CGCGCTGACG CTCTGGAAAA CCTTCCCCAG GGCGATACAG
12481 ACGACGCACG AGGCCAGCTA CCACTTCGTG GCCAACGACG TGACGGCGAC CTTCACGTCC
12541 CCGCTCTCCC AGGTAACTAA CTTCACGGGC ACGTACCCCT GCCTCAATGA TGTTATTCAG
12601 AAAACCCTCA ACGCCACCAT CAAGAAGCTG TCCGATACCC ACGCAACAAA CGGATCGGAG
12661 CAGTACTACG AAACCGAGGG GGGTCTGTTT CTCCTGTGGC AGCCGTTAAC GCCGCTAAGC
12721 CTAGCTGACG AGATGCGCGA ATTAAACGGC ACCACGCCAG CACCCCCCAC CACAACCTCA
12781 ACCGCCAACC GCGTTCGAAG AAGCGTCGGT ACGAACGAGC AGGCAACGGA CGACCTAGCG
12841 GCGCCCCAGC TGCAGTTCGC CTACGACAAG CTCCGCGCGA GCATCAACAA GGTGCTGGAG
12901 GAGCTCTCCA GGGCGTGGTG CCGAGAACAG GTGAGGGACA CCTACATGTG GTACGAACTG
12961 AGCAAGATTA ACCCCACCAG CGTAATGACG GCGATATACG GGCGGCCGGT GTCGGCCAAG
13021 TTCGTGGGCG ACGCCATCTC CGTGACGGAC TGCGTGGCGG TGGACCAGGC GTCCGTCAGC
13081 ATCCACAAGA GCCTCCGCAC GTCCACCCCG GGATCTGCT ACTCGCGCCC CCCGGTCACG
13141 TTCAGGTTCC TCAACAGCAC CACGCTGTTC AAGGGCCAGC TGGGACCCAG AAACGAGATC
13201 ATACTGACGG ACAACCAGGT GGAGGCGTGC AAAGAGACGT GCGAACACTA CTTCATAGCG
13261 AGCAACGTAA CCTACTACTA CAAAGACTAC GTCTTCGTGA AAAAAATTAA CACCTCCGAG
13321 ATATCCACCC TCGGTACGTT CATCGCCCTG AACCTGTCGT TTATAGAGAA CATAGATTTC
13381 AGGGTCATCG AGCTGTACAG CCGCGCGGAG AAAAAGCTGT CCGGGAGCGT TTTCGATATA
13441 GAAACCATGT TCAGGGAATA CAACTACTAC ACGCAACGCC TGGCGGGACT CCGGGAGGAC
13501 CTGGACAACA CGATCGACCT GAACCGCGAC CGCCTGGCCC GCGACCTGTC CGAGATAGTC
13561 GCGGACCTGG GCGATGTCGG CCGCACGGTC GTTAACGTGG CCAGTAGCGT GATAACCCTG
13621 TTCGGATCAA TCGTGAGCGG GTTCATTAAC TTTATAAAGA GTCCGTTCGG GGGCATGCTC
13681 ATGATCCTGG TGATTGTGGC GGTCGTCCTG ATCGTGTTTG CGCTAAACCG GCGCACCAAC
13741 GCCATCGCCC AGGCCCCCAT CAGGATGATC TACCCCGACA TAGACAAAAT GCAGCCCTCT
13801 GGCGGTAAAG TCGACCAGGA GCAGATTAAA AACATTCTCG CCGGCATGCA CCAGCTACAG
13861 CAGGAAGAGC GTAGGCGGTT AGACGAACAG CAGAGGTCAG CGCCCTCGCT TTTCCGGCGC
13921 GCGTCAGACG GACTAAAACG TCGCTTTAGG GGATATAAAC CGCTGGAAAA CGAAGAGGCT
13981 CAAGAGTATG AAATGAGCAA ATAACCACAC CCACACGCCT GTACTTGCCG CCCGCCAGAG
14041 CCGCGCGGCC AATCGATTCG CGCACCCGGC CGGTCCCCGA CACCCTCTAG CGCCCCCGGG
14101 CGTCCGCGGT CTGTATCAAT CATGGATTTC TTTAACCCGT ACCTGGGCCC TCGCGGACCA
14161 CGCCCCCACT CACACAGAGG CACCGATGCT CCCGCCCCTG CCGGCGCCGG AGCCGTTCAG
14221 CCGCCACCAG ACGTTTGCAG GCTCATCCCC GCCTGCCTCC GAACGCCAGG GGCAGGCGGG
14281 ATGATCCCGG TCACGATCCC GTTCCCGCCA ACGTACTTCG AGAACGGTGC TCGCGGAGAC
14341 GTGCTGCTCG CCAACGAACG GTCCATGTGG ACGGCGCGCG ACCGCAAGCC CGTCGCCCCG
14401 GACCCCCAAG ACCAATCCAT CACGTTTCAC GCGTACGACG TCGTTGAAAC AACGTACGCG
14461 GCGGACAGGT GTGCCGAGGT ACCTAGCCGC TTCCAAACGG ACATTATCCC AAGCGGAACC
14521 GTGCTCAAGC TCCTGGGGCG AACCGAGGAC GGCACCAGCG TGTGCGTGAA CGTGTTCCGT
14581 CAACAGGTAT ATTTCTACGC GAAGGTTCCA GCCGGCATTA ACGTCACCCA CATCCTCCAG
14641 CAGGCCCTCA AGAACACAGC CGGCCGTGCC GCGTGCGGCT TCTCGACCAG AAGAGTGAAC
14701 AAAAGAATTC TCAAAACGTA CGACGTCGCG GAGCATCCCG TCACGGAAAT CACGCTATCG
14761 TCCGGTTCCA TGCTCTCGAC CCTCAGCGAC CGCCTCGTCG CGTGCGGGTG CGAGGTGTTC
14821 GAGTCAAACG TGGACGCCGT TCGCCGGTTC GTTCTGGATC ACGGGTTTAC CACGTTCGGG
14881 TGGTACTCGT GCGCGCGCGC CACGCCCCGC CTGGCGGCCA GAGATGCCAG GACGGCCCTG
14941 GAGTTTGACT GCAGCTGGGA GGACCTCAGC GTTCAAGCGG ACCGCAGCGA CTGGCCCCCG
15001 TACCGCATCG TGGCCTTTGA TATCGAGTGC ACTGGAGAGG CGGGATTTCC GTGCGCCACG
15061 CGCGACGGCG ACGCGGTGAT CCAGATCTCC TGCGTCTTCT ACACGACCAG GGAAGGCGCG
15121 CCCAATCCGC CAAACATACT GTTCAGCGTC GGGACGTGCG ACCCCATCCC GGACACCGAC
15181 GTTTTGGAGT TTCCGTCGGA ATATGACATG CTGGTGTCGT TCTTCGCCAT GATCCGCGAC
15241 TTCGAGGTGG ACTTTTTAAC CGGCTATAAC ATCTCAAACT TCGATCTCCC GTACCTAATC
15301 ACGCGAGCGT CCCAGGTGTA CAACCTTCGA TTAAACGAAT ACACAAAAAT AAAAACCGGC
15361 TCCATCTTTG AAGTTCACGA GCCCCGTGGC GGGGGAGGGG GGTTCATGAG GTCGGTCTCA
15421 AAAATTAAAA TAGCGGGCAT CGTCCCCATA GACATGTACC AGGTGTGTCG CGAAAAGCTC
15481 AGCCTCTCCG ACTACAAACT GGACACGGTG GCCAGGCAGT GTCTGGGTGG GAAAAAAGAG
15541 GACGTATCGT ACAAGGACAT TCCCCCTCTG TTTCGCTCAG GTCCGGGCGG CAGGGCTAAG
15601 GTGGGCAGCT ATTGCGTGAT GGACTCGGTC CTGGTGATGG ACCTCTTAAA AATGTTTATG
15661 ATACACGTGG AGATTTCGGA GATAGCCAAG CTGGCCAAGA TTCAGGCCAG GCGCGTCCTG

15721 ACGGACGGCC AACAGCTCCG CGTGTTCTCC TGCCTGCTGG AGGCCGCGGC CAGGGAGAAC
15781 TTTATCCTCC CGGTTCCAAC GCCCGAGGGA CAGGGGGGCT ATCAGGGCGC GACGGTGATC
15841 AACCCCATTC CGGGGTTTTA CGACGAGCCG GTCCTGGTGG TCGATTTTGC CAGCCTGTAC
15901 CCGAGCATCA TCCAGGCGCA CAACCTGTGC TACTCCACCA TGATACACGG ACGAGACCTG
15961 CACCTGCACC CCAACCTGAC GCCGGACGAC TACGAGACGT TCGTGCTGAG CGGCGGACCG
16021 GTACATTTTG TAAAAAAACA CAAGCGGGAG TCTCTGCTGG GAAGACTGCT AACCGTGTGG
16081 TTAGAAAAGC GAAGGGCGAT CCGGCGCACC CTGGCGGCGT GCGATGACCC GTCGCTAAAA
16141 ACCATCTTAG ATAAACAACA GCTGGCCATC AAGGTGACAT GTAACGCGGT TTACGGGTTC
16201 ACCGGGGTGG CCAGCGGCCT CCTCCCATGC ATTAACATAG CGGAAACCGT GACGCTCCGG
16261 GGGCGCACGA TGCTGGAGAT GTCAAAGTCT TACGTGGAGG CCCTGACGAC GGAAGACCTG
16321 CGAACGCGTC TCGGTCGCGA GGTGACCGCC CGTCACGGCG CGCGGTTTCG CGTCGTCTAC
16381 GGTGACACCG ACTCCCTCTT TATCGCGTGC GACGGTTATT CCGCGGAAGC CGTTTCCGCT
16441 TTCTGTGACG ATCTGGCCGC CAGGATCACT GCGGACCTGT TCCCCCCACC CATTAAGCTA
16501 GAGGCGGAAA AGACGTTCAA GTGTCTGCTG CTGCTGACGA AAAAGCGCTA CATCGGGGTC
16561 CTATTGAACG ACAAAATGGT CATGAAAGGG GTCGACCTCA TTCGCAAAAC GGCCTGCAAG
16621 TTTGTCCAGG AGCGATGCCG CGCCATCCTG GACCTGGTGC TCCACGATCC GGAGGTCAAG
16681 GCTGCGGCGC GGCTGTTGTG CAAGCGGCCG CCGCACGCGG TATACGAGGA GGGGCTGCCG
16741 GCTGGCTTTA TAAAAATCGT AGAGGTCCTC AACGCGAGCT ATCTGGACCT CCGAAACAGC
16801 GTCGTGCCCA TCGAGCAGTT AACGTTCTCC ACCGAGCTCA GCCGCCCCGT CTGCGATTAC
16861 AAGACCACCA ACCTGCCCCA CCTGGCGGTG TACCAAAAGC TGGCGAGCAG GTGCGAGGAG
16921 CTGCCCCAGG TGCACGATAG AATCCCCTAC GTGTTCGTTG ACGCGCCCGG GTCCCTAAAG
16981 TCGGACCTGG CCGAACACCC GGATTACGTC AGACAGCACC AGATTCCCGT CGCGGTCGAC
17041 CTATATTTCG ACAAACTGGT GCACGGCGCG GCCAACATCC TCCAGTGTCT GTTCGGCAAC
17101 AACGCGGACA CCACGGTGGC CATCCTCTAC AATTTTCTCA ACGTCCCGTA TAAGCTGTTC
17161 TCGTGAACGC CAATTGGAGA ACGCCAACAT AAGACGCCGC GCCAGCGGAG TCCGCAGGGA
17221 GAGCTCGAGC GGCGAGGAGC GACCAACGGA GACCGCCACC ATGCTGGTTA ACGAACTGTC
17281 GGTGGTCCTC GGCGACTGGG AGGTGACTTT TCACCGGGGT AGATTCAGCT TCGTCAACCT
17341 CACCCGCCTG CAAACGTTCA AGGGCCACGG GGGCTACGCC AGGGTCCGAC TCCCCTTCTC
17401 GCTCGACCAG TTACTCCACC AACATTTCGC GTTCGGACTC GTGACGCGTC TCAAGGAACT
17461 GCCCCCCTTC TCCGACTGCG TGGCCCTTAT CGCCCCGTTG GATTCCGGCG GCGACGCGGA
17521 CGCGGCGCGC GTGGCCCCCG GGTTCGTGCT GGACTCCTCT CGCCCGCTGA CCGTGTGGGT
17581 AAACGCGAGC GGGCGGCACA CGATCCGGTT CTGCCTCCTC TTTCTAAAGC CGATCGACCT
17641 GGAGCGCGCG GTCACGTACG TCTTCGGCGA GAACGGCGGC GCGCGCTCGG AGGGCACCCC
17701 AAAGCCCACC TGCGCGACCG AAAGCCTGCC CGGTGGACCC CTGCGCGTCT CCGGCGAGGC
17761 GTCTCAGACG TCGCCCCATT CTTTCGTTGC GTATTTTCCC ACGGCCAACT CGGTGGCCTG
17821 CCTAAGCCTG TTGCGGTTAC AGGTGAGGCC GTTTTCGGAT GACGCGGCGC ACAGGGACGC
17881 GCGGATCTCC CCGAAATACG TCACGTTTAG TAACTCCGGG GGTAACGTCT GCAAGGCGTC
17941 CGTTCACACG CTGTCCCCGT CGCGGTGTAA AACGGCGCAA ATGGAAATCA TCTACGCTCC
18001 CGGGGACCCC AACGCCGAGA TAGTCCTGGG CCAGTCCGGA CCCGTCCTGC CCACCCACAC
18061 CGGCGGCCGC GTATTGGGGG TCTACGCCGA CGCCGAAAAA ACCATCCAAC CTGGAAGCTC
18121 CGCGGAAGTC CGGGTTCAGT TAATCTTCCA ACAGGGAGCG GCCGCTCGGG GCGATCTGGC
18181 GTTTCTGGTC ACGGGCGTGG CACCGGAGCC CCTATTCGTC GTCACCCCGG CACTCTTGCT
18241 TTCCGGTTGC ACAACCCACC TGCGCCTATT CAACCCCAAC GGTACCCCCA CGACTATAAA
18301 AAGAGACACC CTTGTGGCCG CCGCCGCGCC CTGCCCCGTG GTGCGATTAA GCTCCGCCGA
18361 CGACGCGCCG CGAGACCTCG TCGCGTCACC AGACACCGGG GCGCTCTCCA TTAACGCGTT
18421 CACAATCCCG GTCGGTTTCC CAGGGGTGGT CTCGGCGGAG TGTCACGTGT CCCTACGCGA
18481 CAACGGGGTC CACGAACGCA TGAACCATTG ACGGCAACGA TGGGAACACC GGTGCGTTTC
18541 TTTCGCGGCG AGTGGCAGAC CTCGAGTCTA GTGGACAACG GCACGCCACG GTACAGCTCC
18601 CTGGTGTGGG CCGCCACTAT TCACGACGGC TACCTGACAC TGGTGAACAG GTCAGAGCTG
18661 TGCGTCACGG AGAGGTCTCC GTGTCTGCCG GCATGCCCCA GCATCGGGAG ACTGGTCGGG
18721 AAGAGGTTTC CCGGCTTCGC CTTTGCCAGC GCCACTCTGG GCGATCGGGG AACACGCACC
18781 GTGTTCTACG CGTTCGGTCA CCGCGACAAC CCACTGGACA TAGTACCCGC CGTGGTCGAG
18841 CGCGCGGATC GCGAGCTGGT GCTGCGGGTT CACGCTCCGC AAACAACGCG GGTGTCGCGA
18901 TACGGACTTA AGGTATTCGT GGCGATCGTT ACGGTGGTGC GCCCGCCCGG GGTGTTCCTA
18961 CACTTTCCAC AAGACCGCGT TCCGATCGCG CTGACAGACG CGTGCAGCCA GGAGGGCTCC
19021 AGGCTAACCT CTGAAGAGCC GTGGATAAAA ATTCAAGGCT TTCCCGTCCT ATCTGACGAG
19081 ACCGCGCACC CATTTCTCTT AACCCAGAAG ACCAAGCCCT TTACCGAGCG AAAGTTTTGC

19141 CGCCTGATCA TGGACAACGA CCAGCGCAGC GCCGTCAACA CCGTCTACCT GGGAAAGCAG
19201 CACGTGAGGG TGACCGTGAC CCGCCCCCCG GAAACAATCG TCACCGACGG CCCCGTGACG
19261 GCGACCCTGT CCCTCACCGG TAATGCGCCA ATCGCCTTTC GCCACAACCC ATACTTTGAA
19321 CTCCCGTGGT CGTCCACAAC GGCGATATTC ACGCCCGTGG TGTACGTGGG CCTGACCGTG
19381 TGCATCCCAC CCAACTGTAG CAAATTCGTA AGGTACGGTA ACACCTACGT CTCGGCATTT
19441 AACCGCAAGC TGACGGCGAT TATTAGCAAT CACGCCCACA ACGGCGGGTT CCGGATTCAG
19501 GACTGCGAGT GGCCACCGAA CCGGGAGATA GAGATTTTGG TAACCAACGT GTCCCAGGCC
19561 CCGGTGTACA TCAGCACCGG GACGCAGCTG GGGCAAGCCA TCTTCGTGTT CGCGCCGCGG
19621 TTCGGTGGCC CGGCGAAACT GCGGCAGCTC CTCGGCCACC GATCGCGCGC CCTGGAGCTG
19681 CCGGGCGGGG TGACAGTGGA CAGCCAAAAA CTGTGTAGGT TTGAGACCAT GTACCTGTTT
19741 TCCACGTAAA TTACTAATAA ACCGTTTGCT CGTATCGCTC ACACAACGCC AAACCGTCTC
19801 TCATTCTCGG GGTCGCGCGC CTCGCGAACA CACAAGGTGG CTCAAACACC CCCCTCCCGC
19861 ACCCTCGCCA CACAAAACCA GTTAACGCCT TCCGTTAGAT GCAGTTTATT TATTATTTTA
19921 TTACATCATA GCTATTGCGC GGCGCCCGTC CCGCAAAAAC ATCTGTAGAT ATTCCAGTAT
19981 GCGAAACGCG CTGAGAACAA CGTCCCGGGG GCTCTGCCCT CCCAACGCAC GCACGGTTTT
20041 TTCATTAGAC TCCGCGCCAC CTATCTTGTT ATTTACGGGA AGCTCCTCAA TTAGAGAGTC
20101 GAGGGCGGAC AGCACCACGG TCACCGCGGA GCCAATGGCC GCGGCGTCCG CGTCCCCGGG
20161 CGCGTCGTCC AGCAGCATCC TCAGTCGCGT CAGGTAACTC TCGTAGGCCC GGAGCCCGCG
20221 AACCATGGCG TTCATACACT CGGCCCCGCG AAACTCGCGG CGCTGGCACT CAACGTTAGA
20281 CATCAATATA GGGAAGTGAA CAAACGACAT GATGCGCGGG TACTCACGCA CCCGACAGAG
20341 AGTGGAGTGG TGGCACAGGT AAACGAACCA GCGCTGTATG TTTAACCCCA CTTCAGACAG
20401 GCGCCCCCTG GTCCTGTCAA GAGAGGCGCG GTTGCCGGCC CACTGGGGGA GAACGTTAAT
20461 TCCAGCGGCA GTCGGGGGAG GCGCCAGCGT AGGGCTGGCC GCCCAACACG ACAGGTAAAA
20521 CAAGACGAAC CAGACAGGGA ACATGACCGG TTAAATTACT CGCTTACAAT CGCGGGCGGC
20581 GGCCGGTCAA CGCCAGGTCC ATTAAAAACA CACCGGCCGC CAACCCCCAA CGCGGGGCCG
20641 CGCCCTGGAA CGCGGTTCCT TCCAATCGCA AAGAACCGCG TCACAAAAAG GGCTCGTTTT
20701 GAACCCATTT TGTGCCATCG GGTTTCGTTT TCAGATACGG AAACGGCTCG TCCAAAAACA
20761 CCCAACGGGG GTGTGCTCAA ACGGCCATCT CCATCTCTAT GTGGGGATGG GGGTCGTAGC
20821 CCTCGAGACT CAGATCCGCG CGCGTAAAGT CCTCCAGACG CGCCACCTTT CTCAAAATCT
20881 TCAGCCGCGG AAACGGACGC GGGGTCCTCC GCAGCTGAAG CAGCAGGGGA TCAACGTGGT
20941 TGTTGTAAAC GTGGGCGTCA CCCAAGGTGT GCACAAAGTC TCCCGGGGTC AGGCCCGTGA
21001 CGTGAGCGAT CAGATACGTC AGGAGGGCGT AGCTGGCGAT GTTAAACGGG ACCCCGAGGC
21061 CCATGTCGGC GGACCTCTGG TACAGCTGGC AGGACAGCTC CCCCCGAGCC ACGTAAAACT
21121 GACACAAAAC GTGACAAGGA GGGAGAGCCA TCCGCGCGAG GTCCGCGGGG TTCCACGCGC
21181 ACATAACGAT GCGCCGATCG TGGGGCCGCC TGTTAATTAG ATCCACCACG TAACGCAGCT
21241 GGTCCACCCC CTGACCCTCG TAGTTGGCGT CGGCCCCCCT GTACTCCGCC CCAAAATGTC
21301 TCCACTGGAA CCCGTACACC GGCCCCAGAT CGCCCTCGCG GCGGTCCCCG AAGCCCTGCG
21361 CCGCCAAAAA GGCGCGGGAC CCGTGCGCGT CCCAAATTTT TACGCCGCGG CGCGACAGTT
21421 CGGTGGAGTC GGTGGAGCCC CTGATAAACC ACAGCAACTC CTCCACGACG CCCCTCCAAA
21481 ACACCCTTTT GGTGGTTAAC AGAGGAAACT CGTCCCTGAG GTTATATCGG GCCTGAAGCC
21541 CGAACACGGA CCTGGTGCCC ACGCCCGTCC TGTCCTCCCT CTGCACGCCG TGTTTAATAA
21601 TTAAATCCAG GTGCGCCAAG TACTGCAGCT CGCCGTGGTC GCCGCGGCAC GCGCAGCGGA
21661 CGGCCGCGGC ACCTTCGGGT GCGGAAAAAA CCCGGAGTCG TGCGGCCGCG ATTCCAGCCA
21721 CGCAAACGGG TATAATTTTT TTAACGTAGC AGATACCGAG ATGCACCAGG ACAATCATAA
21781 TTACAACCGA GGTGCTTTAC AACTATAACC GTAAACGGCT GAAGACGAAA CTTATTTTAA
21841 AGGCAGCTTG GGCGGGGCCG ACCACAGGCG CGGCCCGGCT GTTAGCGAGT AATAAACCCA
21901 CGTGGTCTCC CCAAGCGTCT CCCCGAGCGT GAGCGTCTCC CCAAACGCCA AAAACAACAC
21961 AGATAATAAA ATAAATAACA TGTTTATTTT TTATAAACTT AACGCGCGAT TATGCTTCTT
22021 AACAGGGGCC AATAGAAGCA ACGTGCGAGG CGGCATTCGC GCGCTCAAAG GCAACCGCAA
22081 CACCGGCGCT CGTAACAGCG TAAACAAACA ACTATTAGCC ATTCGTAACC GTAAAATCCA
22141 ACCCTCTGCG TCCCGAGCCA TTCGCCCAGC TGCGTCTACG TCGAGGGTGT TTCTAAGTCG
22201 CATTTTGAGC CCTTGGCACG GCCGCGTCTG CGTTAATGGT GTTTCTAAAT CCCGCTGCCA
22261 AGGCCCTCTT CGGAATCGTC AATCAGGCTG CGCCGGGTTC TTCTGCCCTT TTTGGGACGC
22321 TTGTCCACCG CCTGAAGCAG TTTCTTCACT GCGTCGTCCC CGGGATTGGC ACACAGCTTT
22381 TTACCGCGGC GAGTGTTTAA TATCACGGCG TCCACCGAGC ACTGCGACGA GGTGTGGGAG
22441 TAAGAGACCA CTAAACCGGG TGGCGGCAGA TGAGTTACAT ACCCCAAACA GCAGAGTTCG
22501 GGCGCGGGTC CGCTCATGGA GCCCATAGGA AAGGCATAAT CAACTACACA CGCGAACACC

22561 GCAAAAAAAA CGCACACGAA AAGGCCCCTC ATGGTGACAA AAACACAAAA CAAAGACGGG
22621 GTCGCTCGTG CGTTGCCACC AGCGCCACGC TCGCCGTGCT CGGCCACGGT TGACGGCGTC
22681 CATTAATACC CGCGGCCCGG GCCCGCCTCG ACGGAAACCA GGCGTGGTTA GCCACCTGAC
22741 GCACGTGCCT CCCGTCCAGA TGTGGCATTG CGAGTCCCAA AACGGGCGTG TCTAACTACG
22801 GTCGCCCCGG GGTTCAAAGA ACGGGGATCC CCCAAACAGG TAAAAAGCTT TTTGTCCAAA
22861 CAGAAAGCCA ACAGCAACCG AAAGGAATCT CCAAAACAAG CAAGGGTCCC GTGCCGGCGA
22921 TAAACGGAAA ATTTCCCTCA CCTAAACACA CGCGCCGCGT AAACTGTTAA AACACGCAAT
22981 ACTTCCTAGG CCTGTGTTTA ATAAAACACA CAAGGGTTGT AATTTCGCGG GGCCCTTTGG
23041 GCCCCGCGCG TCTGGGGGTG CGTTGTCTCG TTGTGGTTCG CAATCACCAC CCCGCCTGCC
23101 CGAAAACAGG GCAACAACCC CCTGGCCTAG TTTTTTAAAA ACTTAACACC GGCAAGGGGA
23161 GAGGGAGAAG GGGTGCGGCT AAATGGGCTG TTAGCAGCCA ACCGCAGCGG GGAAAAAAGG
23221 GGGCAAACGC CGGGTTAACA CAAACAGCAA CAGGCGCCAA ACCCAATATA ACACAACGCA
23281 TATTCCGTGT GTTTCGCGGG CCGGGGTCGG TACGGCTAAC CCGCGTTGGG CGTTTGCACA
23341 AGCACGCCCC CGCCCTGCAC CCGGCTCCCC GAGGTCACCG AAGGGTCAGG AAACATAGTT
23401 TTCACGGTCG CCAGTCGCGC GCGTGGCAAC ATCTTTCGGT GCCCTCCGAA CGACCCGGAA
23461 AAATCCCCGC CGCGCCGGGT CGTTCGGAGG GCACCGAAAG ATATCCCCAA ACGCAACCTA
23521 AAGCATCATG TTTGGGGTTT CGGTGACGCG CGCCGCAGAG GAGGCCGGTG GTGGCGCTGG
23581 CGAAGATAGT GGCCGCAGAG CACGGGTTTC GGTAGTGACG CGCCTGGACG GCTGCCACCT
23641 GGCAGCAGGC CAAAAACACG GGCGCCGCGA GCGGCCGGCA GGTGGCGCCG CGCTTTCCCC
23701 CACGCCGCGA TTTGGCAAAA TGGCGGACCG GCAGCGCCGG TTGGACGCGG CGGACGTCAA
23761 TCACGACGGA GAGAAGCGCA GTTGGTCAGT TGTCCCGTCA ATCACCGAGG GCCGCGCCGG
23821 TTGGCGGATA ATATGAAACC GGCGCGGCCA TTGGACGCTG GCGGCAGCCA ATGGGCGCGT
23881 TGGGTTTTGT TTACAAGTTC CCTATATATA TTATATTATA CCTTCCCTGT TAAGGGAATC
23941 CCATGTTCTT AAATCTAATC AGCTGGGTTA TATAATTAGA ATATCAGTAA TTACTATTAT
24001 TACTGTATTA TTTATTAAAT TATTTATTTT ATTATTTATT TAATAAATTA ATTCTTGCGT
24061 TATTATGTAA TCAGATACTG GGAAATCTAC TCCCAGTTTA ATAAAAATAA ATAATTAATT
24121 CTATTATGGT GGGGGTTCTA AATATTGTAC AAAACAACAA TATTTTTTAA ATTTATTCAC
24181 TAAAAATGCT CCCATTTTAC TTGTGACGTA CATGTTAGGA ATGTGGGAGT GGTTTGTAAC
24241 TATGTTTTCA AACAACACCC TTTTGGGCCG TCTGAAACTC ATATAAGCGG AAGGCTCGGC
24301 GTTAATTGCC ACTGGCCGCT AACCAGTCCA CCTTGCCAGT TGGAGTTTTA TTGCTGCTTA
24361 TGGCTTTACC TCGCATTCCG GGCACAAGCT GATTTACGTA GCACGCATTG GGGGTTTAAC
24421 CCGGTTGAAT TTAGTGGTTT GCACGTGTAA AAAAGCCCAC CGAGGCTTGG GTTTAAGCGT
24481 TTAAAGCGTT GCTGGTTTTA TAACATGCGA TTTAATTGCT ATGGGAGCTA GCTTGTATTC
24541 CGCATGCTTT CGGTTTAGTA GTTGTACAAA ACAGAATTGT AGCCCGCAAA TTGTTACTGG
24601 TACCCAAAAC GGCAAGCCCG CTTAACACGC CCCGCGCTGC GGCTTTTGTT CCTAGCCCGC
24661 TCGCAATACA ATCGAGATTT TAATTTAAAA TGGGGCAATG GCGTAGCGCA AATTTAAGCG
24721 CGTCAAAAAT TTAAAATGGA CTAACGTAAC CGGTGCTTTT AGTGACCTAT GGCGAATTTT
24781 AAAATTTAAA GCCGTGGGTT TTAACACAGA GCTGCCAGCT TGTAATACGT TGCATATACA
24841 GTGCCGTTGT GATTTTATGG TTAAGTGGAT TTTAACATTG AATTTTGTAA TGGTGTACAC
24901 GTGGATTTTT AAATTGAATT TAAAATTGTA ATTACGTGCG CGTGGATTTT AAAATTGTAT
24961 TTATGTAACA TTGTTAAATT TTAACTGCGT TATGGTTGTA TTTTGGTTAA TTTATGTAAT
25021 TTTGCTTTTA GATTTGTTCG GGTGTGAATT CCAGACAGGT AAGAGCCTAG CTCCTAATGT
25081 TTGCCTTGCC GCCTAGCTCC TAATGTTTGC CTTGCCGCCT AGCTCCTAAT GTTTGCCTTG
25141 CCGCCTAGCT CCTAATGTTT GCCTTGCCGC CTAGCTCCTA ATGTTTGCCT TGCCGCCTAG
25201 CTCCTAATGT TTGCCTTGCC GCCTAGCTCC TAATGTTTGC CTTGCCGCCT AGCTCCTAAT
25261 GTTTGCCTTG CCGCCTAGCT CCTAATGTTT GCCTTGCCGC CTAGCTCCTA ATGTTTGCCT
25321 TGCCGCCTAG CTCCTAATGT TTGCCTTGCC GCCTAGCTCC TAATGTTTAA CAACATTAAT
25381 GTTTAAGCAC ACTAAAATTT AAAGACGTTT GTGTTGGTTT TTATGACCAG CTTGGTACAA
25441 AACCTGCTGG TGATTTTTTA CCCAACAAAT AATAAATAAA AAAGTTAAAA CTTATTTCTG
25501 TTGTCTGTGG TAATTGGGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG
25561 GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT
25621 CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT
25681 CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC
25741 CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG
25801 AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG
25861 GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT
25921 CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT

25981 CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC
26041 CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG
26101 AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG
26161 GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT
26221 CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT
26281 CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC
26341 CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG
26401 AGGGTCCCGG GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG
26461 GCTCCCCGTT CCCCGAGGGT CCCGGGCTCC CCGTTCCCCG AGGGTCCCGG GCTCCCCGTT
26521 CCCCGATTCA ACGTTAAAGT GGTTTATTTA GCAAACACAC GTTTTTTGCC GCGGTTTGGG
26581 TTGTACATAC ACGCGGGTGC GCATTGGCCC GCCCATGCAA AAAAATAACG TTGGGTTTCC
26641 GAAACGCGTA GGTGCTTTAG TGTTTCATGG CACGTTTTCG TGTCACGGTA GGAGTGGTTA
26701 GTTTTTTGGT AGCGAATGCA ACCAGGTAGC TCGCCTGACC GGGAAAGGGA ACGTTTTAAA
26761 CCGCAGTGTT ACCGCGGTGC GGCATACGGT TTAACGTGCA ACTGCCGGTC CCGCGCGTGG
26821 TACCTTAAGT TTACGTGTAG CAATTATGGC GGCTGTTCAG GGCCCTCCGC CGCCCCCTGA
26881 GGAAGAAAAT GAAAATTCTT TGCCAGTTGA CGTTTATGCT ATAGAGGGCA TCTTTCTTTA
26941 CTGTGGGCTC GGGCAGGCGG AGTACTTGCA CCATCCCGTG TTTAGCCCTA TTAAGGAATT
27001 TATCAGCGCC TTTCTCAAGG ACAGCGCTCG CCTGTACGAG AGGCTCTTGC GCCACACCGA
27061 TTACCGCTCT CTGCGGGGAC TAAACGCCAT AGGCCAAGGG ATGCTGCAAA TAAACACGGA
27121 CGGACGCCAC AACTGGGGTC GCGCTTTGGC CGTGTTGGGT CTTGGTGCGT ATGTGGTGGA
27181 TAAGGTTAAA GACGACGAGC GTCTTTTAAC GTTCGCCATA GCCGTTCTAC CCGTGTACGC
27241 GTACGAGGCG CTGGAGTCTC AGTGGTTTCG TTCACACGGC GAATGGGAGG GACTCAGGAA
27301 TTACTGCGAG CGAATACTGA GGCATCGCCG CAACGCGAGG AGACACATGT GCTACGGAGT
27361 TGCGGCTGGT CTTCTGGCGC TAGTGGCGCT GTTTGCCATC AGGCGATAGG TGTGCGTTTA
27421 ACCGCGTATT CCCCCACCCT GACCTTAAGT GAGCCGATTG TATATGAGAC CAATAAAACA
27481 AAACAAGCTA ACGTGGTATT CGTTGGAACA TTTTTTATTT GAGCAGTTCC TCGCAGAACA
27541 TTTTTTGTAT CTGTGACACG GGGGCCGGCT GTGCGCTGGC CTCCACCGGG GCCCGCGCGG
27601 ACTGTCCGCA TTCTGGATCG AGGGGCGGGC ACGCGCCAGC GCGGGGGGCA GAGGCTGAAG
27661 GAATGGCGTT TGACATTACA GATTCTTGGG CGGCGGGGGC TGGAGTTGCC GGCTGCGTGG
27721 CGAGACCGGC GTGGTGGCTC GCCGGCTGCG GGGGCGGTGG TTGCGTTTGC TGGATGCCCT
27781 GCGCGCTCTG GTACGGCTGT GGGGCGTGGA GATGGGTGTC CGTTCCGGAA TACCACTGCG
27841 GCTGGTACGC TAGGTACGGC GCAGACGCGT AGGTCTGTGG CCACGAGAAC TGAACGGGCG
27901 GTGGTTGCGG TCGTTGGTCG GCGTCCTGTG CGCCGGAGGT CTGCGACGCG GCTCGTTTTA
27961 GGTCCCTCAG ATCCGCCTGA ATGTCCAATA TGTTTTTGGA CAGCGCCATG ACGTCCTTGT
28021 GTATGCCGAC CTCCTCGCCT GGAAACACGG GCTCGTCGGA AAGCTCCACG TCCCGTTTGC
28081 GCTTCTGGTG GCGAGACGGT GCCGGAGCGT AGCCGTACGT CGGAGAATAG TACGTGGCCG
28141 CATACGGAGA GGGTTGGGGG GCAAACGCCG ACTGGAAAGG GAACGGTGGC GCGACGGGCG
28201 GTGCGTACGA TGGGACGGTC AGCTCAGGCG GCGGTATCCA CGATGGATAC GCCGCCGTTG
28261 CCGGCATCGA GGGTGGCGCA GAAACGTACG TGCGGTGGCC CTGGTTGCGC ATGTTATCGA
28321 GACTGCTCTG CACCATGGTC AGAAAGGTGC TCTTGGGAAT GGTGATGTTG TCCTCCGGGA
28381 GGGCGCTCAT GGTTCGGGTG TCGCGGTTAC CGCAGTGCTG AGCGCAGGGA AACTGGCTTG
28441 CCTTTAAATA CGTACTTCTC GCGACTCTGG CCACGCCTTT ATCTGTTTTG AGCAGGTCTG
28501 TGCGATCTCG GATAAATCCG GCGTCGATCG CCTTGGCCAT CAGTGTCTCA AGAGGGGCCG
28561 CGAATTCTGG CGGTGTCACC TGACGTTCGC AAAGATCTAA ACAATTGACC GTTATTTTAC
28621 CGGCTTCCTC TCGCGTAAGA GAATCAAACT TGGAGACGAC CCATGTAGGG TCTGGTCCGT
28681 AAACGGCGAT TGACCCCCTA CGCCTCCCCA GTGCGCACAG AGAGACGTGC TGAAACGCGG
28741 GTGGGTGATT GGGGTCTTGC AGCTCTTCTG GGTGGAGGGA CGAGAGGGAG AGCTCCGGGA
28801 GCCATGTGTG TAACATCTCC AACAGGGGGT CCCGAGGCAT GGGGTTTAGC TGGATCTGGG
28861 CCGCTCGCGA GTCCCCCGCG AGCCGTGACG CCAGGGCCAA GAAATCGTGG CTCGTTAGCT
28921 TGCCCAGGCA AAATATCCCG TGTGTTACTT GGAAGAGACC GATTGTGTGA CCTACGTGGG
28981 CTTCGGGGAC GTGCTCTATG TTTATCGGTA GAGGGTCCGT ATACGGGAGC AGGGTCGCCA
29041 CGATTGAGGG CTCTAAATAC AGCTCCTTTT CTATCTTTGG TAGGCTGACC ACGTCCACGT
29101 ATCCCCCAAC GTACACGGGA GTCATGTTGG GTCGGTGTCT GCGTGAGCCG AAGGAGATGA
29161 GCGCCGGTCT TCGGGGGCTG ATGTGGCGCG TTATCCGTTG TGAGAATTTA AATACGTTTC
29221 TCCCTGGCGA GCTCAGGTTT CTTCACCTGG TGCTGTGCGA AATGTACAAC TACGGTCTTA
29281 ACGTGTACCT TCTGAAGGAG GCAATTGCCA ATACCGGAAC GAGGGACGAT ATTGTCCTCG
29341 GGCGTAAGGT TCCGGTGGAG TTCTGGAAAA TCGTTTACGA TGGGTTAAAG GAGATGGGGG

29401 TTTCTGACGC AACGTTGCTC TCAGAAACCA AAAGGGGGGC TCTGTGGCTT TACTTCAACG
29461 GGCGTCCGTG CCTGTTAAAA GGTCTAGGCG ACTACGTTTT TTGCCAACTG GGGTTAAGCC
29521 ACAGCGTACG CGTGGTTCCG GAAAACCTAA CGGATGGGAA TTATCTCTAT AACCTAGGCA
29581 GCGTCATACC ATGTCGCCTG CTTGTGGCAC TGAGCTACTG TTTGGCGTTT TGGGGGCACG
29641 CTGACCACGA GCCCTGGGTC AGGTTGTTCG CCGGCAAGAT TTTTATACTG TACTTAATAA
29701 TTTCCGGCCA CATTATGCCA CGGAAGTCAA TCTTAGAACA GGTGGGAACG TCCGGGTACG
29761 GTGGCTTCGT CGAGGCAGTG TGTCGCGATG TTCGCGCGGT ACACGGGATT CCGGCGTGGG
29821 ATTTTGCGTC CGCGGCTCCG GCGTTAACGT CTCAACAGAC GGACTACTTG TTTGCGTTCA
29881 ATAACAGCGT AGTGTAGGGT CGCCTCACAC TACCGACGTG GCGACCGGCA GGTATCCCAG
29941 GACCAAAAAA TAAACTCTGT CGTAGTCGTC GTAAACTCCG AATTGGCTTT TCATTATCTC
30001 GTGTGCCAGA CCGGGGGGCG ATGGGTGCGA AAGGATGGCG GCCAGGCCTT CCTCGGCCGT
30061 TAAACGTAGC GACGTTTGAG CCTGAATCAG GCTGCCAGCG TCCCTGAACG TGAGCTTTTG
30121 GTTCAGGATG TCAAAGATTT CGTTAAATCT GGAACCGGTC AGGTTGACGA ACGGACGAGT
30181 GTGTTTTAGA TGGGATATGG CCTTTGCCTC CAGCGCCGCT AGATACGCAC CCGGGAACAA
30241 AAACGACGTG GGGGCGTCTG GCGCGCGCGA AAGAATTGGG GAGATGTAAT TTTCTGTCAG
30301 GAAATCAAAA AGTTGTCCCT TTTTAAGGTA TTGTTTGGAC CCGGACGGGT CTTCTACCTT
30361 GGAAAATCTC CCACCCAGGA GCGAGGCCAG GTTGAGGCGC CGGTGGGGCG ATCTGGGTGT
30421 GATTTGGGCC AGCTCCAATA GCACGTAAAC CAAAACTAAA CAGCCTAGTG CCGAGAGGCC
30481 GGCTCGGAGG TTCCATTGGT ACGCTTGAAA GGGAGGCACG TCTTCGCTGA ACAAAGTGGT
30541 CGCAAGGATG TATACGAGAT CTGGTCCAGG TGGCTGACCG GTACCGGGGT TCTTGGAACC
30601 GGGTAGCAGT CCGCTCTGTG CCAACGTCGC GTATATCACG TGGTTTGTAA AGAAGTCCGC
30661 GTTGTACTTT GTTTCTTTGT TTAGGGGCGC TAGAGAGGCG CGTTGCCTCG GATTCGAATA
30721 TGTAAACTCG AAACCCAGGC CCTTCGTGTT TTGTTTTAGG TGAGAGACAA GCGTCGCCAG
30781 CTTTGTGCCG AGGTTATCCC AGAGGCCGTT TCGAGTTGGC GGCAACGGGG TACCTGTCCG
30841 GTGACAGTAG TACCCGTTGA GGAGGCACAG CGCGGCCTGC GTGTCCGATA CGTTGCGGAC
30901 ATCCGCGTAA AAGTCTTGTG TGATCGAGGC GGCTGTTGAG ATTACGGCAG ACATTAGCTT
30961 TAGAGACGTG GAGTTTTGAA AGTTTGTTGT GCCTCGCAAA TCTCTGGGGA ACACCCTGCG
31021 TTGCATGGCG TTATCTGTGA GGCTACTGTA CCACGGCCCG TATTGTGGCA GCCATTGGTT
31081 TTGGTTCATG TACAGGCTTG GCAGGAGCTC CAGACGGAGA TCCTGGTTTA CCGTGAATCC
31141 CGGGTCTCCG GGGGCTATTG TGACAACGAA CGCGTCTCCA TGTGTCTGCG GGGCGATTTT
31201 GGGTTGCGGG TTTAGTGATG TTTGGGCATT GGCGGTTTCC CTGCCCGGGG ACCGCAGATC
31261 CCCGACCATG TTTTGAATTG TTTCTAGGTC GCGCATGACC ACTTCTGTTT CCCGTTGGTG
31321 CGTTTGCAGC ACGTTGTCTA GTTCTAGCCG GAGCAGGTTT TTTCGCAGCG TGTGAACCCT
31381 TAGCGCCAGC TGTTGATTTC GCATCACAAA CGCGTTTCTT CTTTCTTCTA TCAGAGATGG
31441 TTTGTGAATA AACAGAACGT TGTTTGGATG GGGGGCCCAA AAAGTTGCGG TAATTTGCCG
31501 CGCTGGTTTT CGTGGGTATC TCATACAGCA TTTTTCTGAC GTGCGCATCT TCCCATTGGC
31561 TAAAAAACCC GTTAAGCGCC GCCGACGTGG TGTTGATAAA TGTAGGCGGA AACGCTGATG
31621 ACTCAGTATG AATTGTTTTA AGTGTTTTCT GGTTCTTAAA TAGCAGCCAG GGCCTTACCG
31681 ACCAGTTTTC GTGTCCTCTT GGCGCGGCTC TGCCTATATA TGCCACGGAA TCTGTCAGTT
31741 GGTGTAAGCC CTGCCATCGC TGTGATATTT TTATTGCGTT GCCGGGAAAT ACAGCCGCGC
31801 TTAGGCACGT CTTGAGTTCA ATTACGTAGC ACGTTTTTTG ATGTTCGCTA TTAACGGTAA
31861 GCACACAATC TGCTATTCGC CTACCTAGGG AAACCTCGAA AAAAATCCTA AATTTAACGT
31921 TTTTGGTTGC CCTGAGTGTT GGGCGAATTG AAAGAAACTT TAAAATATCA TCAAGCGACC
31981 GGTGTTTGAT TAACCGTCGG TATACTGCTA AGTGGGCGCG AGTTCCAGCC CGTTTGCGTG
32041 AGGCCGGAAG GGCTTCGAGG GTCGCGACGT GCTTACATTG ATTTGCAAAC GCCATGGCCG
32101 AAGGAGGGTC TGGCTTTGGC GACGAGCTTG TCCGTCAGAT GCGGGATCGG AAGCCTAGAT
32161 GGGACGAGTC TTCCGACGAT ACTGATGACG TTGACACCGA AAGCACCGAT CTTGAGTATG
32221 ATGATGTGTT CCCCGTGGTA GATACGCACG GCTTAATGAG CCCTGGAAGT CAAAACTATG
32281 ACGTACCCAC GTCTCCGTCC GGAACGCCGT GGGAATTACT GCACCCGGAC GCCCTGTATG
32341 CACATCCGAG GTGCCCGCCT AAAAGGGCGG TGGTGCCGGG CGGTGGCGCG CGTCCCAAGG
32401 TGTCCGCGTT CTCGGCTAGA CTTCAATATG TTGGACGACA GAGTTTTGGG GATAGAGAGA
32461 CGCGGCAGCT TACCGGGGCT CAGTTTTCTT CCGAGAGCGA GCACGAATAT GCGGAAATAC
32521 CGGAGCGTAC TACTACACGC CCGGTTGAAA GCGGTGACAA GAGAAATTTC ACCTCCGGGC
32581 GTAGGGGTGC AATCTCGGGA CCCTCGTCGA CAAAACCTAG TCACGGTGCG GGGTTAACAC
32641 GTAAGACTAA AACGTCGCTG AGTGTTAGTC TTAAAAACTT GCTGCGGATA AAGGACGATG
32701 ATGCTAAGGT TGATGTACCC AGGCCGGTTA CCGTGCCGGT TCATCTCATG CAGCCGCATC
32761 CTATGACGGA ATATAGAAAT GCCTTTTTAA TCTATCTAGA GGGGGTCATG GGTGTTGGAA

32821 AAACAACGCT ATTGAATTCA ATGACTGGCA TGGTGCCGCA GGAAAACGTC TTAAGCTGTC
32881 CCGAGCCCAT GAAATTTTGG ACGTGTGTTT ATTCAAATTG CCTTAAAGAA CAGCGCAGCA
32941 TAGTTAAGCA AGGCACCCAC GGGAAATTGA TCACTTCTGC TCGCGTATAC GCGTGTCAGA
33001 GCAAGTTTGC GCTACCGTTT CGTGCGACAG CCGCCGGCAT CGGTCGCAAC CTGCAACCGT
33061 GGCTGGTGGG AAACGGTAGC ACAAAGCCAG CGAATTGGAT TGTTTTTGAC AGACACCTAC
33121 TGTCCGCAAC GGTGGTTTTT CCGCTGGTTC ACGTAAAGTA CAACAGGCTA ACGCCGGATC
33181 ATCTGTTTCA AATCCTATCT CTTTTCTCGG CACACGACGG CGATGTGGTC GTTTTGCTAA
33241 CGCTCAACAG TTCGGAGGCG CACAGGCGCA TTCAAAGTCG AGGCCGTAAG GAAGAAAAAG
33301 GAATCACGCA AAACTACTTG CGACAGGTAG CGTGGGCGTA CCATGCCGTG TTCTGTACGT
33361 GGGTGATGAT GCAATATCTC ACACCGGAGC AAATGGTTCA GCTGTGTGTA CAAACTGTGT
33421 CCATAGAGGA CATATGCAAT ATGAATTCCA GATTGACTCA TCGGTTTCTA ACCCTAACAA
33481 AACTACATGA ACAGAGTATG ATACCCATGG TGGCAGAAAT GTTAGTATCG GTTAAAGAAC
33541 ACGTGACCTT AATGGAGGTC TGTTTGGGAC TCTTTAAAGA GCTACGAAAG CTTCAAATTT
33601 TAATTGTTGA CGCAGGAGAA CATCTAGATG ATGCGTGTGG CCTCTGGGGA AATATTTATG
33661 GGCAGGTGAT GTCAAATGAG GCTATTAAAC CACGAGCAGT GAACTGGCCA GCTCTTGAAA
33721 GCTACATTCA AACGCTAACC AAATTGGAAG GCAATGGCGC GTATTAGTTT TATATTTTTT
33781 TTTACTATAA TAAGGTGTTC TGTAACGGAC AAATACGTGT ATGATGAAAA ATCAAATGTA
33841 GAGTTGGAAT TTAATGGAAC AATTTACCAA ATTAATTGGA GAAACGTAAG CAAAGAATTA
33901 ACATCTATTG TAATGGAAGA TGCGTGGTAC GATTCTTTAC TGCTGGAACC ATTGTCTGTG
33961 ACTTTGGAAA AAAGAAAAAG TTTGCTGCGA TCAAGTATAG TTAATGTGCA TAATAATGAC
34021 TATACATTTT GTAAATCTTC AAGTGACCAT GTAATTAACC TTACGGTGGA TTTTAATTAC
34081 AGTTCTTTGC CAGGTTTTAC TGGGAATTTT AATGTAATGA CACATGCTTT AACACAAGGC
34141 GTTTTATTAA CGAAACGTGA ATTATTTACC AACTCAACAA ACATTATGGA TCTTTTTTAC
34201 GCGGAAAAAA TAAATGCAGA AATGTTCAAG ATTACTTTTG ATTATTCCAA TGTAATAATT
34261 TCCGGCATTA TAACAGAAAA TTGGATATTA GTATCAGTTA CAAATTCAAG TGTAAAATCA
34321 AATATGCAAT GTGTTGCACT GTTGTTTGGT GTTCCTAGCA CATTTCCAGC GTTAAAAGGC
34381 TATGTATCTT ATCGTGATTT ATTGGTCGTT AAAAATTCAA ACTATGCTCT GGGTGTAATT
34441 GCGCCGAAAT CTTATAATAC TCTTGATTTG GCATTTTTAC CAAAAAATTT TACAGAAATG
34501 TTTGTTTCAG TTATAGATAG TCCGCTTAAT GCCATTGATT ATCTCAAAGG AAAACTTTTG
34561 GCTATTGAAG CAAAGGGTGC CTGTCAAAAT CCTTCAAATG AAAATGATAT CTTATCTTTC
34621 TTTTTTGAAG TTACTGCAGT AAATTTTTTA TTTATAAAAA ATCTTCAAAA GCAGCAATTA
34681 GTTAACGTTG GGTGTGTTGT ACGGCATGTT GCAGCGTTAG AAAGTTTAAT GCATTTATTA
34741 AGGCTCTGCT ACCCAACGTT TAAGTTATAT GAGCTAAATT TGGAAACGTT GTCGCATATT
34801 GCAGAGTCCC AGGTTTTTAA CTTGCCGGCA AATAGCATGT TATCGTTATC AGTAAATGAT
34861 CAAGAGGTCG TATTTTCAAT GTTTAAAATT GTATATAATA CACCTAAGGT TGGTGGTAAA
34921 ATTCTAAACG AAATTGTATA TATTACTAAT TACATGTACA CCAAATACAG TGAAAATTAT
34981 CAGCTAACCA ACACATTTAG GCGGAATGTA ATGAACATGT ATGAAGTTTT AACAACAATT
35041 AAATTAAACG TCACAGATTC TTCTGTGTTT TATCCGTATA TATTGTTCAC ATCCATGTGT
35101 AATAATGTAG AGATATCATA TATGATTAAT CAAATAGCAA AGCCCGATGA CATCACCATA
35161 TTTCGCGTGT TTTCGCCTTG TTTTTTAAGC TTAAGGTTTG ATTTAGATGA AAACAAGCTA
35221 CGATCCGACG CACCGCAGAC GTCAAAAAGA ACCGGCTCGG AACTGGCACA AGGAGCGTCT
35281 GGTTTTTGGC GGCTGTTGCA CGCTTTTCAC GCCACGCGTA TAAACGAGTT TTCAGTCATA
35341 AATTGTACCA GGCTGGCATG GAAGCAAGTA ACCGCGCTGA TGCCGCTTAC GAACATCACG
35401 TACGTCATAA GCTCCGTGCG CCCTGATCAC GCTCGCGTTT ACGAGGTTTC GGAGGTGTTC
35461 CTTAACAGCG CGATGTTTGT CTCGGCCGTT TATCCAAATT GTTCTCATTT TACACCGCCC
35521 GGTACCGCCC TACACATTCC GATATTGTAT AACTTTTCCG CCCCGAGGAT AGGATGTCCG
35581 CTGTGTGATT CGATTGTTTT GAGCTATGAT GAGAACCAGG GACTACAGAC CATGATGTAC
35641 GTGTCAAACC CCACAGTTCA AGCCAACCTG TTTTCCCCGT ATTCTCCTTT CTTTGACAAC
35701 GATAATTTTC ATATCCATTA TTTGTGGTTA ATGAACAACG GTACCGTTGT GGAAATTCGG
35761 GGATTGTACA GAAGACACGC ACTCAGTGCT ATTGCATTAG TATTTGCTTT TATTGGAACA
35821 ATGTCTGCTC TGTATTTTCT GTTTAAGCTG TTTTCCATCC TGGCCTAAAC TGTTAATAAA
35881 GCGTAAAACT TTAAAAGGTT GTTTCCCGTT TCTTTTTGGA TGGTTACGCG GGGACTGAGA
35941 TTAGCGAGAA CCTCGGATAG AGGCGTGCCG TCACATATTG ATTCAATCAC CTCTGCGCCA
36001 ACTTGAATGG CCAGCTCTCG ATTTTGGGGA GGAAACATTA AAAAATATTT GTCGGTCGCG
36061 GTTAATAATC GCTGCGGAGC CCGCGACAAA ACTCGAAAAT CCTTCTGAAA GTCGTGGGAC
36121 TCTATGGAGT TTAAACAAAA CACCATGTCT TCCAGGCGCT GTTTTACGCG CGAGAGCATG
36181 GATTCGAAGT ATTCTCTCTG ATACGGGCCC GCCGCCTTAC ACGTATTCGC GTATAGCATT

36241 TTAAAGTGAA ACAGAATATT TTCCAAACCC AAACGCTGCA ATAGGACGTG CTTTAAAAAA
36301 AGGTGTTGGA TAGGCTTTAG TTTTAAATTA GGTAAAAACA TTATAGAGTT GTTGGCGCGC
36361 TGCATAATTG TCACGGTGTT AAACGCGGTA ATCGGGCTCT CGCCTATTTG TGGCTGGCGC
36421 GAACCCATGG CAAAAAATAT GTCTGCCAGT AAACAGTCCG ACTGTACATT TGACTCGCTG
36481 GTCGCTGGCG GTTGCTTGAG GATGGATCGA AGTTTTTTAT TATCCGGCGG AAGCGCGCTA
36541 CATAAGGTCG GTGGCGCATT TGGATCGTAT TCTTGCGTGG TGTTTGCCAG CTCGTTTATG
36601 GTTATCTGTT TGTTATTTGT AAGAGTTACC GGTGGGGTCT TGACGGTCGC CGGGACGGAA
36661 AATGGCCTGC AGTGCAAAAA AAGAGAGTTA CTCGCAATCA CTAATTTTGA GTTTACGGTA
36721 ATTGGTGTGT CCTCAAAGTC AGAACTTCTA ATTACTGGGT CTCTGAACAA TGATACGCCT
36781 CTCTCGCCTT CGTCCGTTCC CACGCACAAG CAGCTGTTTA GCGCGTAGGG ATGGGGTGAC
36841 ACCATAATTC GCAGCATTGA TAATTCCGGA CGTTTTGGGG CGTACAATGG CATCATGTTT
36901 GAATAGATGG TGCTGAATTT TTTAAACAAC AGGGGAAGGG GTACGGCCGG TCCCGGAAGT
36961 CCGACCATTG CGGCAAGTGC CGACGAGTAC ACGACATGAC AGACGTACTC GCCGTTGGAT
37021 AGCTGCACCG CCCCCCCAAC TAAGCGAGCC TTTAAATCTG AGATTTTTAT CATTGATTTT
37081 CCTTTGGTAG TTTCAGGAAC CGCAGCGGCC TTGGCTTCGC GGTTATATTT TTGCTTTATG
37141 TCTCCAGTGA AGTCCCAAAC GTCTAATAGT GTGTCGAATT ACCTGCGGTT TTAAGTTTGA
37201 AAACGCGTGA CCGGTCGTAG CTATGAAAGA CATAAAATTG GCGTTGCGTC CTAACTCGTC
37261 TAACCGTATT TTTTTACTTC GCACGTATAT CTTTGGAATG AGGTTGCATT TTTTATATCC
37321 CAGAATAGCG CATTCGTGTT TGCTTTTGGT GTGACTAATG ATGCGTTTTG AAAAGTCAAA
37381 CGTACCAAGG CGGGCGGCGC GGTTCAGGGC TCCCAGATCC AAACGCAAAA ACGAACAACA
37441 TTCCACCAAA CACTCGATAA AACGCTTGTG TAGAGTTTGT ATTTGCGCTT TATTTCTTGT
37501 CTCCAGCGTT GTGAATATTG TGTTTTTCAC CTGGTCTATT CTAGCATATG GAACGCTTAA
37561 TATTGGGTTC AATTCATCCG AAAGTTTTTT AAGTAAGGGC CTACCGTATA TTTTAGAAAT
37621 TGTCGCAGTT GCCGCCTTAA ACCAGGAAAC GTTGCATCCG TCTTTACTAG AAACCACAGA
37681 AGGGAAATTA GTGGTCCAGA AAACGTCTCG CCGCCCCCGT CTAACTATGT ATTGGTTTTC
37741 AAGCCCGTGG AGGTCTAACG TAGAGTTCCA GTTCGCAATC GAAACCGGGA AAATGTGCTT
37801 CACCGGTAAA AAGCACCTGA GGTTTCCCAT AGCGAACATA AAGTGCAGCA TGTCCCCGCT
37861 TATGTTAACG TTAATTGCAG CCCCGTTGCA CAGGTTGTCG GAGTAGAACA CACTCACGGT
37921 GAATGACGGT TCTTTCACGG AGAATTTTTT CATTACAAAG TTATTAGTGA GCCTGGGGAT
37981 GTCGATAATT GACTCGAAGG TACCGGCGAG AGACCAGCAC GTGATTTTGG TATTAATTAC
38041 CATATTGGTA TTGAAAACTG AGATTTTGAA GTTGTGTAAA CTCATGTGCT TAACTGGATC
38101 GAACATGTCT GGAAGAGGTT CGGTCTGGTG GTCGTGACGC GCTTGTTCAT ATCGATCTGG
38161 GTTTTGATCG GCGTTAAAAA TAAAATTGTA ATCTCTATCT TTTGGTGCGT ATTGAATTCG
38221 CGAAAACATG GCCTCTAGGA CGCGAGGCGT TTCGCCTGGC CCCAGAAACT CGGTGTTAGG
38281 ATACAGCCTG TCGCACGAAT GGCACACGTG TTCCGCCACA GACATGCACA TTGCGACGCG
38341 GAATATTGGT ACCTCGGGTG GAAACGTTAG ATAGTACGTT GACTTTGGCG TCACAACAAC
38401 TACCTTGTAC GGTTTTTTGC CGGGATCCTT TTTGTTAATT TTCACCTGTA GATTTACGTC
38461 AAGCGATTGA TTTACTATGT TTTGGTAGTC TATGATTAAT TCGCTCGCCA CTTTGATGTC
38521 CTGAAATTTG TTTTGCTTTG AGAGGTTTTT TAAAAGTACC GGATTGGATT CGATGTTGTG
38581 AATAATTACC GTTTGGACGC GATGCAGTTC GGGATGCGAA ATCCAATCCG AAGTAAAACC
38641 CAAAATGTTA AATTTTAAGT AATTCAGATG CAGTTTTGCA CACGGGGGCG GACACCGAGC
38701 TAGCGTGAGG ATTTTGAATG GCTCTAAAAG TTTAAAGTAA TTCTGATAAC ACGTGCCGAT
38761 CAGATGTTCC GGCAGTGTGC GCGTAAAGGC AGCAAAGCAT TCCTCGAATT CCCATATTTG
38821 AAGGTAACGA CAGAGGATTC GAAACATGCT TTGATAATTT ACGCCACATA CCATTAATGG
38881 CGAACAGTAC AGCGACCGTA GGATTCGTTC GTCGCTTTGA ACCAAAACCG TGTCGGACAG
38941 TTCGTAGGTG AAAACGGTGG GTAAAAACGT ATGCGTCTCG ATATGGGAAA ATGTACATCC
39001 GATAGCATCA TCAATTATTA AACCCGGTCC CGCGTACTCG TTACCGGTCT TGCCAACGAT
39061 ACCACCTTGG CGGTGTATAA AAGGGTATAG TCCTTCTCCC AATACCAGAC GCCGGAGCGC
39121 ACGGCACGCG TAAAAGGTGG TCTCTGTTTG CATAGCCCTG TCCAAGGCCT CTAGCGAGAC
39181 GTTTTCAAGA ATCTCATTAT TGGTGGCCAG GCAGCAAAAA ATTGCCAGTT GAGTTTTAAA
39241 ATCGGAATTG GCCGGTGACG GGGCGTTCGC TGCCACCGTA GCAGGACACG CGGGCAGAAG
39301 AACAGGGCCC TGCAACAGCA TAATGGAGGC GGCTCTGGAG GTACGACCCT TTCCATATAT
39361 GGCCACGGAG GCCAATCTGC TTCGTCAAAT GAAAGAGTCG GCAGCCAGCG GACTTTTTAA
39421 AAGTTTTCAA CTGCTGTTGG GAAAAGACGC CAGGGAGGGA GGAGTTCAAT TTGAGGGTCT
39481 TTTGGGCGTG TATACCAACG TAATTCAGTT TGTGAAGTTT CTGGAAACGT CGCTAGCGGT
39541 TGCGTGCGTT AACACCGAGT TTAAGGACCT CAAACGAATG ACGGATGGAA AAATTCAGTT
39601 TAAGGTATCT GTACCGACCA TCGCGTATGG GGACGGCAGG CGGCCCACAA AACAAAAACA

```
39661 ATACATTATC ATGAAGGCCT GCAATAAGCA TCACATTGGT GCCGAGATAG AGCTGTCGAC
39721 TGATGACATC GAGCTGCTAT TCATTGACAG AGAAACCCCA CTCGATTACA CAGAATACGC
39781 CGGGGCCGTA AAGACGATTA CCGCCTCTCT CCAGTTTGGC GTGGACGCGC TGGAGAGGGG
39841 CCTGGTAGAT ACCGTATTGA ATGTTAAGCT TAGGTCCGCC CCGCCGATGT TTATTCTAAA
39901 AACACTATCA GACCCGGTCT ACACCGAACG GGGTCTAAAG AAGGCTGTTA AGTCAGACAT
39961 GGTGTCCATG TTCAAAAGCT ACCTTATGGA TAACTCGTTT TTCCTCGACA AATCAGACAT
40021 CGCCGTCAAG GGAAAGCAGT ACGTGCTGTC GGTTCTCTCC GACATGGTGG GGGCGGTGTG
40081 TCACGAAACG GTTTTTAAGG GGACGAATAC GTATCTGTCT GCATCGGGAG AGCCAATTGC
40141 CGGAGTCATG GAGACCACGG AAAACGTAAT GCGAAAACTG TTAAACATGC TAGGTCAGGT
40201 TGACGGGGGC ATGTCCGGTC CGGCGTCTTA CGCCAATTAC GTTGTCAGGG GCGAAAATCT
40261 CGTAACCGCC GTGACGTACG GTCGCGTCAT GCGAACGTTT GACCAGTTCA TGAAACGCAT
40321 AGTGGACCGT CCCAACGCGC AGCCCAGCGT CGACGATGAT CGGGACGCGG TGGCGGACGG
40381 GCAGGACTCC CTCGCCAAAA CACCGATCGC GGCGGCTGTT ATCCAGATTG GAGATAAACT
40441 GGTGGCTTTG GAAAGTTTAC AGCGAATGTA CAACGAGACC CAGTTTCCCT TCCCGTTAAA
40501 CAGGCGTATG CACTATACTT ATTTTTTCCC CATTGGGTTG CACATGCCCC GTCCTCAGTA
40561 TTCCACGTCG GCCACAATTA AGGGGGTCGA GCACCCAGCG GAACAATCCG TCGAAACGTG
40621 GATTGTAAAT AAAAACAACG TACTGCTGAG TTTCAATTAT CAAAACGCGC TCAAGTCCAT
40681 CTGTCATCCG CGCATGCACA ACCCGATGCC GTGTGGGCAA GCTTTGGGTC AGGCGTTTCC
40741 AGATCCCGGA CACGTTCACA GGTACGGACA AAGATCTGAG CACCCCCCGA ACATGAACCT
40801 ATACGGGCTG GTGTACAACT ATTACCAGGG CAAAAACGTG GCACACGTTC CGGATGTCGC
40861 CCTAAAGGCG ACGATGACCA CAGATGAGTT GCTGCACCCA ACCTCGCACG AGACGCTTCG
40921 TTTGGAGGTT CACCCGATGT TTGATTTTTT CGTTCATCAG CAGCCTGGTG CGCAAGCCGC
40981 GTATAGGGCG ACCCACAGGA CTATGGTGGG TAACATTCCA CAACCCCTGG CGCCCAACGA
41041 GTTTCAAAAC AGCAGAGGCC TGCAGTTTGA CAGAGCGGCG GCCGTGGCTC ACGTGCTGGA
41101 CCAGTCAACG ATGGAAATTA TCCAAGATAC GGCGTTTGAC ACGTCGTACC CACTACTCTG
41161 TTATGTCATC GAATGCCTCA TTCACGGACA GGAAGACAAA TTTTTGATTA ATTCTCCTTT
41221 AATTGCATTA ACCATTGAAA CTTACTGGAA CAATGCCGGA AAACTGGCGT TTATTAACAG
41281 CTTCCCTATG CTGCGATTTA TCTGCGTTCA CCTGGGCAAC GGTAGTATTT CTAAGGACGT
41341 GTACGCCCAT TACCGAAAAG TTTTTGGCGA ACTCGTGGTT TTGCAGCAGG CGCTCTCGAA
41401 AATCGCGGGC CACGAGGTGG TGGGGCGCAG GCCCGCGTCC GAGCTGATTA ACTGTCTTCA
41461 GGACCCCAAT CTTTTGCCGC CCTTTGCTTA CAATGACGTT TTTACCAACC TGCTCAGGCA
41521 GTCCTCGCGG CACCCCATGG TACTCATAGG CGACGAGGGG TACGAAACGG AAAATGACAG
41581 GGATACGTAC ATCAACGTCA GAGGAAAAAT GGAGGACCTA GTCGGTGACA TGGTTAACAT
41641 TTACGAGACC AGAAACAACG CGGATCATGA CGGCCGCCAC GTCCTTGACG TCGGTCCCTT
41701 TAATGAAAAC GAACAGCACA TGGCTGTGCT GGAAAAGCTT TTTTATTACG TGGTTCTGCC
41761 AGCCTGTACC AACGGTCACG TCTGCGGCAT GGGCGTCGAT TTTGACAACG TGGCCCTGGC
41821 CCTGACGTAC AACGGCCCAG TGTTTGCTGA CGTCGTGAAC CCCGACGATG AGATTTTGGA
41881 CCACCTGGAG AACGGGACGC TCCGCGAGAT GCTCGAGGCT TCGGATATAC ACCCCACCGT
41941 TGACATGATT CGAACTCTTT GCACGTCGTT TCTCACCTGC CCGTTTGTTA CCCAGGCCTC
42001 CCGTGTTGTG ACTCAGCGGG ACCCCGCGCA ACTGTTGACC ACTCACGACG ACGGGAGATA
42061 CGTGAGCCAG ACTGTCCTCG TTAACGGGTT CGCGGCGTTT GCTATCGCAG ATAGGTCTCG
42121 TGACGTTGCC GAGACCATGT TTTACCCGGT GCCGTTCACC AAGCTGTACA GCGATCCCCT
42181 GGTGGCGGCC ACGCTCCACC CGCTGGTCGC AAATTACGTG ACGCGCCTGC CGGCCCAGCG
42241 CGTGCCGGTC GCGTTTAACG TCCCCCCGGC CCTCATGGCC GAGTACGAGG AGTGGCACAA
42301 GTCTCCAATG CTGGCCTACG CTAACACCTG CCCGATGACG CCCACGTCGT TGAGCACCCT
42361 GGCGAGCATG CACATGAAGC TGTCCGCGCC GGGGTTCATC TGCCACGCAA AGCACAAGAT
42421 TCACCCGGGC TTTGCGATGA CCGCCGTCCG AACCGATGAG GTGTTGGCGG AGAACTTGCT
42481 ATTTAGTGCC AGGGCCTCGA CGTCCATGTT TTTAGGGCAG CCATCGGTTA TGCGTCGGGA
42541 AGTCAGGGCG GACGCAGTCA CGTTTGAGGT GAATCATGAG TTGGCATCGC TGGACATGGC
42601 GCTCGGTTAT TCTTCCACCA TCACGCCCGC CCACGTTGCG GCGATTACCT CGGACATGGG
42661 CGTTCACTGT CAGGACATGT TTCTCATGTT TCCCGGGGAC TCGTACCAGG ACAGGACCCT
42721 CAACGACTAC GTTAAACAAA AAGCCGGATG CCAACGATTC GGTGGTCCTG GCCAGATTCG
42781 TGAGCCCGTC GCTTACGTTG CGGGGGTGCC GCACTCGGAC AACATACCGG GTCTCAGCCA
42841 CGGACAGCTG GCCACGTGTG AGATTGTTTT GACGCCCGTT ACTGCAGACG TTACCTATTT
42901 TCAAACCCCC AACAGTCCCC GGGGACGGGC ATCCTGCGTG ATCTCGTGTG ACGCGTACAA
42961 CAACGAAAGC GCGGAACGTT TGCTCTTTGA CCACTCCATC CCGGATTCTG CCTACGAATA
43021 CCGCACTACG GTTAACCCAT GGGCGTCGCA GCAGGGCTCC CTCGGAGACG TGCTGTACAA
```

```
43081 CTCAACCTCG CGCCAGGTCG CAGTGCCAGG GATGTACAGT CCGTGTCGCC AGTTTTTCCA
43141 CAAGGACGCT ATTTTGCGTA ACAATCGGGG CCTGAACACA CTAGTCACGG AATACGCGGC
43201 CCGCCTCACG GGAACGCCGG CGACCAGCGC GACGGACCTG CAGTACGTGG TGGTCAACGG
43261 AACGGATGTG TTTCTAGAAC AACCGTGCCA GTTTCTACAA GAAGCGTTTC CCACGCTCGC
43321 CGCCAGTCAC AGGTCTCTGC TGGACGAATA TATGTCGAAT AAGCTCACGC ACGCCCCTGT
43381 GCACATGGGA CATTATATGA TTGAGGAAGT GGCCCCTATG AAAAGACTAT TAAAGATCGG
43441 AAACAAGGTC GCCTATTAGT TTAGCTCAGA CGGTCTGGAG CTAACGAGAG ATGGCCCTCG
43501 ATAAGAGCAT CGTTGTCTCG GTGACGTCTA GATTATTCGC CGACGAGATA GCAAATCTTC
43561 AGTCAAAGAT AGGATGCATT TTGCCTCTCA GAGACGCCCA CCGTCTGCAG AATATACAGG
43621 CGCTGGGTCT GGGGAACCTG TGCTCTAGGG ATTCCGCGGT GGATTTTATT CAGGCATATC
43681 ACTATTTGGA CAAATGCACT CTCGCCGTGT TGGAAGAGGT CGGTCCCAAC AGTTTACGGC
43741 TAACGCGCAT TGATCCCATG GACAATTATC AAATAAAAAA CGCGTACCAA CCGGCCTTCC
43801 ATTGGGATAA CTACTCAGAA TTGGTAGTTA TACCACCGGT CTTTGGGCGC AAAGATGCGA
43861 CCGTCTCACT GGAGTCTAAC GGGTTTGATG TGGTTTTCCC TGCCGTGGTG CCAGAACCAC
43921 TGGCTCAAAC AGTGCTTCAG AAGCTGCTGC TGTATAACAT ATACTACAGA GTGGCGGAGA
43981 CGACGCCCAC CGACGTCAAC CTAGCCGAGG TGACGCTGTA CACGACCAAT ATCACTTACA
44041 TGGGTCGCAA CTACGCCCTG GACGTGGACC CCGTTGGGTC GAGCTCAGCT ATGCGGATGC
44101 TGGACGACCT GTCCATTTAC CTGTGCGTTT TGTCCGCGTT AATTCCGCGC GGGTGCGTAA
44161 GGCTACTGAC CTCATTGGTG CGCCACAACA AACACGAATT AGTCGAGATT TTCGAGGGGG
44221 TGGTGCCACC TGAGGTACAG GCCCTGGATC TCAACAACGT AAGCGTGGCC GACGACATAA
44281 CGCGCATGGG TGCCCTCATA ACCTATCTAC GAAGTCTCAG TTCTATATTT AATCTGGGCC
44341 GCAGATTTCA CGTTTACGCG TTCTCATCGG ACACGAATAC CGCTTCCTGT TGGTGTGCAT
44401 ATAACTAGAA ACGGGCCTCC CTGTGCTTTG ACATGTCGAT CCCCAAAATT ATGACGGTGT
44461 CCAGAGACAA CGAGGGTACG GTGTGTGAAG TCGCGGTGGA CAACGGACGA CACAGAGCGA
44521 TGATTTATTA CCCTAAGACC ACCAACTTAG CAAACGAGCG CGCGGACGTT GTTAAGGAAG
44581 CTTTTGATAC CGAAACCCCA GTGGACATTG TAAAGCAAAT TGTTAACGAG GGCCTAGCTA
44641 TATCCAAAAA AAATTGCGTC CGTTTGGCGT TGTATTTATA TTTTTATTTG CAGTACGTGT
44701 GCTTTGCTCT GCTCCTCACT TGGCAGTTAA ACCCGTACAT GGACCCACCG GGTCTGGTGT
44761 TTGCGGTTAA CCCCATGGGT CCAAAACATG TCACGAAACT ACCGCACCCG GCTATTGTTG
44821 CGGTAGGTTG TGGGGCAGAC GCCATCTGTA AGAACTGTAG CGTCCCCGAT ATCAAAACGG
44881 AGCTTGGAAT GGTTTACCAC AACGGGTCTA GCGATTCTGG TCAGCGCGCA CACTATGGGC
44941 TGGCCCTGTT AAAGGCGGCC TGGCTTGTCA TGGGAAATGT GTGTCCGGAA CCAGTAGTGC
45001 GGCAAGGCGC TGCATTACTT GGTCCATGGA ACCGGACGGA GTGGTCGGAT TTTAAATCGG
45061 CAATGGCGGC AACCACGTTT TGCGGATCCA GAGGCGTTCT GTGGTCACCG ATTCATGAAA
45121 AAAACCTCTG TCGCCCCACC TGGAATGATG TAATTAACAC ATCAGTTTTT ACAAATGAAT
45181 CACTCTGTCC AAATATACCT GTGGTGCCCG AAAGTGTAAT AGTGCTTAAT GGTGATGCAT
45241 GAGAACAATA AACGTATTCC CACGCACTTC ATGTACGTTG TTTTTATTGG TCTGGTAAGG
45301 TATTAACAGT AATGGGAGGT TCCGCTGGTG CCTATAAAGC AAAACGGTCA TAGAGTAACA
45361 ATATTGTGAA TGGGGAAAAT AGTCTGTAAA TAGTTTATCC GAACAGTATG ACTGCACACA
45421 CGAATGGGGT TTTAACCACG ACGGGCTTTT CAACAAGTCA GCCGGAATCG GTTCAAGTTT
45481 CTCCATTTTA TCGCGTAATT ACAAAACCTC CCGTTATGGG CTTGTTTTTT TGCGTGGCTA
45541 TGTGCGTTAT CGCGTTGGTA TGGTACGTGA TGCGGAGGGT GTGTTGTAAG GGGCGCGTTG
45601 TTGCCGATTC GTGTCGCGAC CCGCGTCAAC CCGCGTATGA GATGTTGAAT GTTAGGTTGC
45661 GTCCCCACGG AACCAATCCA TAGAGAACTT TCACGTACAT AGCCAAAGCG CACGACGTGC
45721 TCTCTAGCTC GTTTATTCTA GCGATTTAAA AACAAATTTA TCGTTGGTTG CCATGTAGTG
45781 GGCCATGACG GCGGCGACAA GCACGTCGTC CGACATGGTT TTTTGTTTCG CGCAGTACGT
45841 GTGACCTCCG TCTTTTAGTG GAATGCAACG TATTGCCTTG ATTTGATCTA TTAGATACGC
45901 TACGGGATCA AACGATAGCT TAATAGTGTG AGACACCACA GTTTGGCTAG CGCTAAACGT
45961 TCCCGAGTTC AATGCGTAAA TAAAAGATTC AAACGCCTTG GCTTTCTCTG GCCCCAACAT
46021 GTAAATGGGC GAACGTATAA GAGTGTTCTT GTCCACGTGG TGTAAAAAAC TAAGAGGGAC
46081 CGAGCAAATC TCGTTTAACA CGGTTGCTAT GGCCACTCCG GCATCTTGAC TGCTGTTGCC
46141 TTCCACGGCG ACGTTAACGT GCGTGATCTG CGGGTGAAGG GTAACGATCG CTCGAATTAA
46201 CGCGGCGGCA CATGACGCTA TCTGGTACGC GGCGGTGCCG GTTAGATCTC TTAGAAAAAA
46261 ATGTTCCACG CCTAATAAAA TGCATTTAAT CACCTTGTGA TTTACCGCAA TCACCGCACC
46321 GATGCCGGTG CCGGACGCGT CTGTGTTGTT GGTGTACGCC GGGTCTATGT AAACGTGGAG
46381 ACACGGCTTC ATCGCGCCCT GTATATCTTG TGACGTGCTC TTTACGCGGC ACAGATCCAA
46441 TTGTGACAGC GAAGAGTCGC TAACGATTTT ATGCATGCTT TGAGCGGACG TGGCAGCGTC
```

46501 CCCCATTAGC TCCGTGGAGA ATGCGCCTTC CAAAAAAAGA TTAGTCGTGC TCCGTACCGT
46561 CTCATCAATC GTAATATAGG CAGGAATATG CAGGCGGTAG CACGGGCACG CCACGACAGT
46621 GTCTTGCAAG TTAAAATCAT CTTTATGATC CGGACATACG TAGTTCACCA CGTTAAGCAT
46681 CTTTTCGTGG GCGTCCTTCA AATTAAGTAA AAAACTAGTT GATTTGTCAG ATGAGTTACT
46741 TGAGGATATA AAAATCAGTT TTGCGTCCTT TTGAAGCATA AAACCCAGAA TCGCGGGGAG
46801 GGCATCTTTT TTAATGAAGT TCGCCTCGTC GATATATAGG AGGTGGAACG TCTGTCCCCG
46861 GATGCTCTAA AGGCAAAGAA AAACACAAAA CGAGTTACGC GGGTATGGAG AACGATACGC
46921 CTAAGGACAA AATCTCGGAA GCTGACTTTC AACAGTGTCA GGCGTTCTTT CACCGTCCCA
46981 TTAGAGATCT AATTTCATCT GGAGCTGACG CTTTAAACCA CTTTAGCCTA TCTGAATCAG
47041 ACGGACATAA ATTGGAACGG ATTGTTCTTC TGCTTGACCT GGTGGGGACA GAATGTCTCT
47101 CTTATACCAC GATCGCTGCA AAGAATGTCA AATGACGCGC GTCAACAGCC CAATATGTCG
47161 ATTTCATAAC GTCTCTAACT TATACCAGTG TTTGGATTGT AAGCGCTATC ACGTATGCGA
47221 CGGGGGACGC AACTGCGTGA TCGTGTACAC TCGCGAAAAT CTAGTGTGTG ATTTAACGGG
47281 AAACTGCGTT TTGGATAATG TGCAGGACGT ATGTTCGTAC GGTCCTCCAG AACGCCGCGT
47341 ACCCGACGCC TTCATCGATC CGCTCGTGTC ACACGGCACG AGGGAATGTC TTAAAAGCGA
47401 TATACTGAGG TACTTTGAGA CGGTCGGTGT GAAATCTGAG GCATATTCTA CCGTTGTCAA
47461 GAATGGACAA TTGAATGGCA TCATAGGTAG ATTAATAGAC GCTACGTTTA ACGAGTGCCT
47521 TCCGGTAATG AGCGACGGCG AAGGTGGCAG AGACCTCGCG GCGAGCATTT ACATCCACAT
47581 AATTATCTCC ATATACTCCA CTAAAACGGT ATATGATAAT CTTCTATTTA AATGTACGAG
47641 AAATAAAAAA TACGACCACA TTGTAAAAAC TATCAGAGCG CAATGGATGC GCATGGTCTC
47701 AACCGGCGAT CCGTCGCGGG TCAGTGCGAC GGGTTGTTTC ACGTGATACT TCCGCGAGGG
47761 TTTATCCTCG CGAACAATAT TACGTGCGGT GAACGGCAGC GGTTTTTTGC ACACACTTGG
47821 TTCGCTGCAT CTGGACGCAC GTCTAAGACT TTATACGTGT GGGGACGGGT ATTTCAAAAC
47881 ACCGACCCGG GCCGCGGGGA CGGTCCGTCC GGGCCGTGGT CCGGACTGGC GATTAGTCTG
47941 CCTCTGTTTA CCACAAATGG AAAATTTCAT CCGTTTGATG TAGTTATACT CAAGGCCGAT
48001 ACGCCTGACT CTGGAAGCTC GTGGACCGTG AAGTTCTTGT ATATGTCATT AATTGCGGCT
48061 TACAGAAACG CAATGCGAGG TTTAAAAGAT AAAGTTTCGC AATGTACCGA TGCCGCCGTT
48121 GACGGTGAGG TTCATCCTCT AACCGTCTTA AAAGAAGCGT TGGTATCACC GGACACTGCT
48181 ACGCGACCCG TGTCCGCGTG CAACCCTCTA CAGATGTTGA CCGGACTCTT ACAGTCTAGG
48241 GTACGGGACG ACTACGTGAC ACACCACCGT GCGCTCGAAC GCCCAGGTAA TGTGAGGGGA
48301 CAAGTAATCG CCCCGACGCG CACCGAGATG CCAAACGGAT CGCCAAGTCG TGTAAGGCTT
48361 GGATTCCGCC CTCCCAAACA AGCCAACTAT CCAAAGACGT GGGCGCAGGC GCGTCACGTT
48421 TTCTCGTCTC GCGCATATTA CGTGTGCGTA TATGATAACG AAGAACTAGA TACCAAGTGG
48481 CAGCGGCAAG ATCCGCGGCC GTTGCCGCTA GATTGGTCCG ATCCGGTCGC GTACCTGTTA
48541 GAGGGCGATT TGTTTTTAGG AGCCAAGCAG AATGCGTTTG TCGATTCTAT AGAAAAGACG
48601 TGCAGGTGTC AGAACTATAC CATTAAGCAA TTTTTTCCGG TTTTGATAAA TAGGGACAAC
48661 GAAACAGTCG ACTTAATTAA GGAGCATTTT ATAGAGGCGT GCTTCGTGAT TAGAAACCAG
48721 GTGTCAGAGA GGAGCGCTTG GGTAAAGGCG GCGCTGTTTC GCAACGATAG TAACACGTAT
48781 TGGAAGGATG TTTTGGGATT ATGGGAGCAT GGGCCTCATA AGCTGGGTAC GGCTATAAAA
48841 CTACCAACAT CGGAACCTTG CAATGCCGAC GTAAACTGGA GCTGGCTCCT GTGTGACGAG
48901 GATATAACTC GGTCAATTAG CGGGCAGTCT ACTGTTTGCT TAGTTGTCTC GCCTACCCTG
48961 ACCGCCTGGC TGGTGCTCCC GGGGGGCTTT GTTATTAAAG GCCGCTACGA CCTATCAAGC
49021 GAGGATTTAA TGTTTGTGGC TTCGAGATAT GGCCACCCAG CGTCGTCACA TTCTTAAATC
49081 GTTTTTAAAC AAGGAATGCA TATGGTTGCG ACACCCGGGT ACGTCCGCGT TTGTTCGGGT
49141 ATACACCGCG ACCACTGCGC ATTCTGCCGT TTTTGACCCG CCGGTAACTA GCGAAAATGC
49201 GATGTCACTT AACTTTTTAA ATGTTATGAT CGTAATCATG AAACCAAAAG AATTTGGCCC
49261 GTGCGTTACC GTGTACATGA ATGGAGATAT TCTAGATTTT TGTGCCACGG AATCTGTCGC
49321 CATAAGGGAC GTGCCTGGTA GGGCGGACCT GTGTTTAATT CGTTTTGGTA CCCTTTCTAA
49381 TGCGCCGAGG AGCGTTCCGA TACCCGGGCC GTTGAACCCA CATCCGCGAG AAACCGTGCC
49441 CGGGCTAACA AAACAGGAAA TTATATACAC TTCGCAAACA GTGCCAAGAG GACAGATACC
49501 AGATGCCATA AAGGGGAAAG AGTTCCACCA AATAAATCCG TTTTTGTGGT TTGACGGAGG
49561 GGCGTTTTGG CAACTGTTCC TCTCTGTGGA TTTTATGCTG CTCTGTCCCG CACTCGACAC
49621 AGTTCCGTCC CTGGCCAGAA TCGTTGGGCT TCTTACACAG TGCGATAAGA GCACGTGTAA
49681 AATTTGTACG GGGGCCCACG TACACGTTAA CCCGTATCGC GGATACACGC CACCTGACTC
49741 GCAAGGGACC TCACCCTCGT GCCCCTGCCT TATCTCGTGC GGGGCCAGGC GCGCGGCGGA
49801 TGTCCTGGTT ACCGGACACG TTAATCTTTT GGGCCTGCTC TTTGACCCCA AAGCCTCCCC
49861 CAAAGTGACC AAGCTGCGTT TAAAAAGAAA CCCACGCCCG GTACCGATAG AGGACGCCAT

49921 GTCGGGCGTC ACGGCCGAAG GGACCGAGGT GCAACCCACT TCGCTACCGT GGGCCCTCAT
49981 TCGCCTGCCG GATTTAGCCA GTCGCGTGAT GCTATACGGC TGCCAGAACT TAAAAAGCAT
50041 CTGCTTACGT TCTTATTGAA GCACGTCGCA CACATCAGCG AACTCGGCCG CCGTCCGGGG
50101 CTCGCGTACA CGATGGTTCC GTTTTCCTTC TTTATGTTCA GGTTCTTCGC CGGAAACCAC
50161 CTGGATAGCG TGGCGATGAC CTCGGAGAAA ACGGCATTGG CGACGTGTTT TTGGTGGGCC
50221 ACGTAACCCA CGTGAACGTT TTCAACCGAC GAGAGAAGAA CGCTGATAAT GGCGACCACG
50281 ATCCATGTTT TCCCGTGACG ACGCGGGATA ACAAAAACGC TGGCTTTCTG CTTAAACGTC
50341 TGTAACACGT CGTCCGTGGT TTCAAATAAA CCAAAGTGTT GCTTAAACGT GGCAAATAAC
50401 TGGTGTGTCT TTTCGGGTGC TTTGATAGAC GCGATAAAAT AAAACGTGTG CATGATTAAT
50461 TGTTGTTGAA AGGGCTCAAG GCAACAAACC CCGGGCACGT AACTCCCATT TAAAAATGAC
50521 GACAGGTTGG CTAAAAACTG CCGCAGTTCT GCATAAACGG GGCACTCTAG GAAAGCTTCG
50581 TGGGTCCGTT GGGCCGATTG GTATTCCATG TACTGGTCCT TGGCGTTTGC CATTTGGACC
50641 CGACAACATG TTTGTTTAAG GTAGTTTGTG AGTTCCTCGG TTAACCGGGG CAACGTAGCG
50701 TGGCTAGAAA TTCGGGGCTT TCCCAACGTA GGGGCCGGCA TGGGTTCTTG AACTCTGCGG
50761 TACGTTTGAA GATAATCGTC TAAAACAGAA CTATATGCAT TAACCGCGTG AATAACTCCC
50821 AAGCATGGGT GGGCCATTCG TTCGGTCTTG TCGTTGCCCG ATATAATAAC GGGCGGTTGC
50881 CGAAACCAGT TTTCGCAACC ACCGTCCGTG ACCACGCGCA AGTTATTTTG AAGACGTTCG
50941 CGATAGCTGG TTAACAACAT GTTCCCTTCA AGTTTTTTGA ATAACGGGCA CCCCGAAACG
51001 GAACGCCGCT TCGTTAAAGG CGTTCAGTTA GCTTTAGACC TGTGTGACAA CACTCCGGGA
51061 CAGTTTAAAC TAGTTGAAAC ACCTCTTAAT AGTTTTCTTC TGGTATCCAA CGTTCTGCCG
51121 GAATCGCGCC CGGTTAGAGA CTGTCCGCAG CCGGAAGGGT TTGACTTTGA ACACATTCAC
51181 CTCCCAAAAC TAACACGCAT GCAGCGTGTC CTGGGGCGAT ACTGCGACCA TGTTAACAAC
51241 GACGACACGT GCGTTAACGT AAAGGCAAGT TCCTCGAATT CACAGGGTGC CTTGTTTTAT
51301 CTGCCGTATG GACAGGACGA ATGGAATTGG GCGCTCACGT TAAGGAAAGA CAAGTTGGTT
51361 AAAATGGCTG TAGAGGGCTT GTCAAATCCC ACGACCTGGA AAGGTTTAGA GCCCGTGGAT
51421 CCTTTACCGC TCATATGGCT TCTGTTTTAC GGTTCCCGGT CGTTCTGTCG GGAACCAGAG
51481 TGCCTATATG AACGCAATTT TGGTATGAAG GGACCCATAC TCTTACCGCC ACATATGTAT
51541 GCCCCCCAAA AGGACGTAAT GACTTTTGTC CATCATGTAA TTAAGTACGT TAAATTTTTA
51601 TACGTGAACG CCGGTGGGGG TCTTGAAACT GAACCGTCCC CGCCGTTCGA GGCCTCGCGG
51661 TTGCGCGCAG CCATCGCTCG TCTCGGGGAC GTGGAAGCGG ATGACGCATA CCTGTCCGCA
51721 AAGTGCATGT TGTGTCACCT GTACAAGCAA AACGATACGA TTTCGATTCA TGAAACACAC
51781 GTGGGCGGAG TCATCGCATT AGGCGGAGAC GGTGCGAGAT ATATAACGTC TAGTGTTCGG
51841 GCTCAACGGT GCACGAGTCG GGGAGATTTT GTTTTAATCC CCCTGTACAA CATTGAAGGG
51901 CTCGTAAGCA TGATAAGGGA ACATGGCCTC GGCAGCAGCT AAAAAAATGT TAATTAAGTC
51961 GGAGCTCGAG TCGGAAATCA ACAAAAAACT GTCCATCTCC GTATTTGACA GGTTTGGGGC
52021 CGACAGTGCC GTGTTTAACG CGCAGTATAA GGGAACCAGG GAATCGCTGC GGTCGTACAA
52081 CAGCCTAAAA AAGAAGGACG ATCTGGCGAC CGTTGTCGGA ACGCTAGAAA CGTCGCTGCG
52141 TGAAAAACAA AGCGAATTGG GATTACTAAA GGGGTTTAAC AGGAAAAAAA TTGAAGAGTT
52201 TGACGCTGTG GCGGACGCGG TTCGCGACCT CAAGGACGAG CTGTACGGAG AACTGGAGAT
52261 TCTAGGTACG CTTGACAATG AATCTGTTCC CGTGGAAGAA GAGTCCCCCA AGGACGACAT
52321 TATTAGGTGG AAATTGGAGC GTCTGCCCAG AGTGTGCCCC AAAAGCCCTT GATCCCATTC
52381 CCAAGGTTCA GACTGACGTC GACAGAACAG CATCGTCCCA TATAACCGTC ATTAAAACAC
52441 GTAAGACGAT CGCCCAACTG AAGATACCTA ACAACTGGGG CCAGTGTAGT CACCAGGCGA
52501 CGGACTGGAC CGCCGTGCTC GGACGCGGCT CGTATGGTGT GGTGAGGTCC ATGTCTCTCG
52561 GCCGCTGCGT TAAGCATTTT GGCAGCCGGC GTGAGTTTTT TTACGAGTGC ATTTTTAACG
52621 ATATAGTACG CGCCTGCCGG GAGAAACATC CCCTGAACCG CGGGGGTGAC CGTATACTAT
52681 GTTTCCTAGA GCCGTGCGTA CCATGTCGCG CCCTGATATT CCCGCAGTTA ACAGGGAATC
52741 TGCTAAACGC GGATCTTAAA CACGTGAACC CTGAACGGCT GGCCGTTGAA TTCTCTGAGC
52801 TGAGGGAAGG CGTTAGTTTT CTAAATAATA TATGTGGCAT CGTTCACTGT GACATCAGTC
52861 CAGAAAATAT ACTGATAAAG GGGGAACTGA CAACTGCGTA CGGGAGACTT ATGATCGGAG
52921 ATCTAGGGTC CGCCTCTTTA CACACGGGAA CCCCTTGGAC CGGAGTGATG GTGACCTCCA
52981 AACTCGGGTT CGTGCAGCAC ACGTACCATT TTAAGGCACC GGCCAGATTT ATCTGTAAGC
53041 ACATTTACCG GCCGTCGTGT CTCCTCTACC GGTGTCTGCT GTCGTGCGCC GGGGGCCCGC
53101 AGGCGCATAT GCTAAATCAG CCGTTCCAAA TCACTCCACA ACTCGGTCTC ACAATTGACA
53161 TATCGTCCCT GGGTTATAGT TTGCTAGCAT GCCTAGAGAA ATATCTTCAG CCAGCTGACC
53221 CATTTCCCCA GCAGGGAGCG TTGGCGGACG CTTCCTCCGA ATCCGCCCAC CCATTGTTCT
53281 ATTTGCGTTG CATGGTGCCA AGAGTAGTCA TCGCCGAGAT TTTTTCTGTT GCCTGGGACG

53341 TTCCACTCGA TTTAGGCATT GACTCATCTG GCCACGCGCC AGCTATTCCC CTGAGAGAAG
53401 CGTACAGGCG GTTTTTTGCC AACCAGTGTA GTTTATATAG GGCGCAATAC AAAGAGGATG
53461 CGTTAGAAAA CGCATCCTCG CGGCTGTGTA ACTCAAAACT TAAACTAGTT CTCCAGAAGC
53521 TGTTGGTCAG GGACTACTTT AGTCATTGCG GAAACTGCGG AGATCATGGA TTTTTTCTCA
53581 GATGAGCCGA TGGTTCAGGA GATGGCGCTT CTCGACATCG ATGAGCAGCA GCGGCTCCTC
53641 TCGAAAATGA GCCTGGCCAA CTTTTTAAAA CATGAGCGAG TAAGGGCGTT TTTTAGCGAT
53701 AACAAAAAGG AAATAAGCAT GCCGGCTATA CGGTTCGTGT ATAATTTTTA TCTATTCGCC
53761 AAGGTGGGAG ATTTTATCGG CAACACCGAC GTGTACGATT TTTACGTCAC TTGCGTGTTC
53821 AGGGGGAGGC GCCTGACGCG CCTGTCGGAA GTGTACGACG CGTGCCTAAA CATGCACCCG
53881 CACGATCGAC ACCACGTGTG TGCATTGATA GAACAGGTCA CGCGCGGCCA AAACATCAAT
53941 CCTCTGTGGG ACGCTCTGAG GGACGGCATA ATTTCGTCGT CAAAATTTCA CTGGGCCATA
54001 AAACAACAGA ATTCGTCCAA AAAAATTTTT AACCCGTGGC CTATAGTCAA CAATCACTTT
54061 GTAGCGGGCC CGCTCGCGTT TGGACTGCGT TGCGAGGAAG TGGTTAAAAA GATACTGGCG
54121 ACGTTGCTGC ATCCAGGCGA GGCGCACTGT GAAAACTACG GATTCATGCA GAGTCCTCTC
54181 AACGGGGTTT TTGGCGTCTC CTTGGATTTT GGAATTAACG TCAGGTCTGA CCCAAAAGAC
54241 GGTTTGGAGT TTCACCCAGA CTGCAAAATC TATGAAATAA AATGCCGGTT TAAGTACACT
54301 TTTTCCAAGA TGGAGTGTGA CCCGATTTAC GCTGCGTATG CTAAACTTTA TCAGAAGCCC
54361 AGCATGCAGA CGCTTAAGGG GTTTTTGTAC TCCATATCTA AACCGGCGAT CGAGTTTGTC
54421 GGAGAGGACA GGCTCCCCAG TGAATCGGAC TATCTTGTGG CATATGACAA AGAATGGGAG
54481 GTGTGTCCGC GGAAAAAGAG ACGCTTAACT GCAGTACACC ATCTAGTTAA AAAGTGCATG
54541 ATTCACAACT CTACGGCGCC TTCTGATGTG TATATATTGT CAGATCCGCA GGAAACCGGA
54601 GGCCAAATTA ATATTAAAGC TCATCTGAGC GCCAACCTAT TTATAAACGT CAGGCATCCG
54661 TATTATTATC AAGTGTTGCT CCAGTCTCTC GTTGTACAGG AGTACATCAG TCTCTCCAAG
54721 GGAACTAAAA ATTTGGGAAC CCAGAAAAAC TTTATAGCGA CTGGCTTTTT TAGAAAACGG
54781 CAGTTTCAAG ACCCGAGCTG CTGCACGATC GGTGAATTTG CCCCGTTGGA TCCACACGTA
54841 GAGATACCGA CCCTTTTAAT CGTGACACCG GTGTATTTTC CCAGCGTGGC CAAACACCAA
54901 CTGGTGAAGC AGGCGACCGA ATTCTGGGCG GCTAGTGCTC GTGAGGCATT TCCCGAGCTA
54961 CCATGGGATT TATCCTCTCT GTGTGCAAAC GCCCCACCAA CACCGTAGAT GTGAAGGGGG
55021 AGCCCATAGA TGTATCCAAA GAATTCGATC CTATTATAGG AGAAGAAAGC ATTGTCTTGT
55081 TAACGGCAGA TGGGACTGCC CCCGCGGCGC TGTACAAACC CAAAACCAAG CCATCCAAAC
55141 ATAAAAACAA TAAATTGTCA GATTTTGTTT AAGCATTCGT ATCTTTATTG GAGTGGGGTG
55201 GATGGTGTGG GGTTGGGAAG GGAATGGGAT TGAGGGGAGG ATGAAATGCT AGAATCATAT
55261 GTATTTTGGA TATGCATCCT CGTAATCGCT CACGTCTTCT TCTGTTTCAT ATAACGCGTT
55321 CAGTTCCGGG GAGTACTTAA CTTTTGTCGT CTTCTCTTTA GGGCCTTGAG GAAGGGGTGC
55381 GTATTCGGGC GACTTTGCGC ACCATTTCCA ATCGTTTCTG ATAATTCTTA AGATGATGGC
55441 TAGGATGCAA GTGATCGGAA TCATAGAGAT GTTCACGGCT ATGACTTTGT GCAGGCTGGC
55501 AGATATGAAT ACCTTGTCGT ATCGAACGAC GGGAAGAAGC AGAATCAGGT AACTGACAAA
55561 AACGCCAACG TAGAATCCGA TCTGGTGCTT TAGGTACTTT ACTAGAAATA GTTCTGTGTT
55621 GATATACCAA GCTATTGTAA GGACAAAAAA CAAGTTTATT GAACCAAAGA CAATATCGGA
55681 AACGAGCATA TAGAAACTGT TTCCAATAGC CATCATCGTT CCGAGGGAGA ATACTAACAT
55741 CTCCATGGCT ATTAAGGACA GATAAAGATT GGCGCCAATC GGTTTTCCGT AACGCAACAG
55801 GGTGTCTAGC AAGCTATTCT CAGGAATCTG CTGCTCTAAG ACGCGTAGAG ACCAGGTGGC
55861 CGAATTGCAC GATATCATAG CCTGGACGTG CGTAAACGAT AGGCAAAAAT GTATACAGTA
55921 TACAAACGAG GCTAAAAGTA TATGTTTGTA CGACAGAACG TGAATAAATA ACTGGATGGT
55981 CCACAGCCTC AAAATAGACA TAAAAACTAC ACTGTGGGAG CCTATGAGCA CGATCCATGT
56041 CTGGAGGCTG GTCATTGTTG TCGCGTGCAC GCGCTTGGCC TTGAATATGG CCAGGATGGC
56101 CCACGCATAG TAACACAGCA GATATCCGTC CACCAAAAAC GCCCACGAGA TGTAGACAAA
56161 CATTTCTGGT GCTTCCAAAA ACAAAGTCGG CGTCAGATGT TTAGCAGAAC TTCTGACCGT
56221 TAGGTTTAAC GCGCTGTAGT TGACCAACGT GTTAAAGTAG CACGGAAATC CAAGTCCCGG
56281 GAATGTGGCG GTCAGTGGCA CTACCGCTGA CATTATAAAC ATAAGTCCAA GAATAACCAG
56341 TAGTTTTACC CATGACGATA AGATAAACGA GTCGCTCCGT GAAATTTTCA TAGTTCAAAG
56401 AATCTGCTCG CGGGTGCGTT CGGCAAACTA GCAGTAAAAT CAGCAAGTGG CTTACATGTT
56461 ACGGATTTAA ATAGGCTCCG CTTTTAAATA TCACATTCTG TGCAAAGGAG TTGAGCCACG
56521 GCACCATGAA CGCCCGGGAG GTGGCACTCA CGGGACATGT TTTGCACATA TCGCTGCATA
56581 GCACGCACGA GCGCGAGAAA TTAATAATCT GGCAGGTTCA TTTACTTGTA TGTCAACAAT
56641 GCGGAATTCA GGGAGATGCC GCATATCTAT TTGTCACCGA AACATTAAGC AATACTGACT
56701 GGGGAAATAT ACCGGCGATA AACCGTCACG CACCGTCCAT AAATGAGCAC GGCCGTAATT

56761 ATATGCAGTG GGAACTCCGT ACTCGTTTAC GGAATCCCAT CATTCAATTG TTAAGTCGCC
56821 AGCCCGGTGC GGTTAACGTA AGGGTCAGCG AGCCGAATAT GGTAATAGTG GGCTGCGAAC
56881 GAGCGTTGGA TCACTCGTGT TCGGTGCGCG TGACTGGAGC CTATCTTCAT TGCGATACCA
56941 CTATGGACTT TAGTTTGGAT TCTGTTGTGT CCCCAACCCG CGAATTTTGG TTCTCAGAGA
57001 TGTTTTCTCA CTGTTTAGTT TCCAACATTG AAGTTTACCT TAAAACAACG GGCGGGTTAT
57061 ACTATAGGGC ATCGAGTGCC ACGCAATGCC GAAAAAGGGC GAAAGATGGC GCATTGGGTA
57121 TTCTTGATAT CTTTAATTGC GAATCTCGTG AAATACAAGT TGCCGGGCAG AAGTACACCT
57181 TGAGTATCGC CACCGCAACA TTTCACGTTC TCTGGGTGGA CGAGGCGTGT ATGTGGAACG
57241 GGGCCCTGGC CGAATTTTTT AGGGCGCTGC ACAATAAGTT GTTCGGCGAC CGGGAAGGCG
57301 TAGCGCCAAC GTTAACGTAC GTGTGTCCGG GGGCCACTCC GGAGGGAACC CCCTTCCCCC
57361 CCTACTTTTC CGCGTTTCCA CACCTCCCGC TCGTGTTTGG AAGACCGCGA AGGCTCGACG
57421 TAACCGCGGT CCAAGAACTC CCAAAAGCAC AAATTGCGGT ACACTGGCCC CCGTTTAAAG
57481 ATTCAATCTT AGGGGATCAG CTTCTCATAC CTGGCATTTC ACCTAAAAAG CCAGGTACCG
57541 TACCCGTTCG TTGGCCGCTT TGGGTGGAGG ATGTTAACTT GAGTCTCTGC GAGACGACAG
57601 AAAGCGTTGC CCGCATAGTC GACCCACATT CTATAGTAAT CATAAAATTT TCATCACTGT
57661 TGTGCCAGCA CCTAAAATGC CACCGTGCGT TTGTTAAAAA TGAGTTAGAA TACATAGCAA
57721 CCATCTGTTC CAGCGACCTT CGCCTCTTCA TCCAAGAGGA ATACAACCGG TTACTTGCCA
57781 CCATTTTTAC GTGGGCCGCG GCGAGCGGGT ATACCTGGGC GGCCATTGAT AAAACAACAG
57841 TATTCATCAA GGCTCCCCAG CTCAGCGCAG CTGTAAGTGG TTTCTGCCCA TCACTAAATA
57901 GTTGCCGTAG GAAACAATGT TACGAAGGTT AAAAATAACA GTTCATTTCC TTTCACAGGA
57961 ACAGCAAAAG GTCGTGACCC GTCTTGAGGC GCATTTGGGA CTTCCCGTAC AGGAAACTTC
58021 CCACCCGCCT GACTGGCTCA AGTGTGAGGT CTGCTCCGCG TCCGTGTTTT TAAAAATACC
58081 AGCCGGGGTT TTGTATGCCG GACTCGCAAG AGACCCCACC AGGGAAGCAA AACGGGACTC
58141 GTGGCTGGAC TGTCTAGTAG AAGGCGCGAC GTTGTTGCTT AACAACTCAG TGTTACCGAT
58201 TGGGGCGCTG GCGGGTATCT TACCCACCCT TTTTGCCAAC AGGCGGTGTG TTAATTTTTG
58261 GCTGCTGCCA CGCGCGTGGG TAAAATCGGC GCCCATATGC CCTCCCCTAC CGATTGACTG
58321 TGTTACGCCT CCACAGTTTG TCGTGACAAA GCGTGGACCA ATCTGCTGGT ACAAGGAATG
58381 GCCGTTACCG GTTGACGTTG ATTTTATGTA CTACCTACAG GAGGCACTAT GTGTTTTTAG
58441 TGTTGTGTCC AACGGGGAGG GTACGGAGAG TCACGCGGAC AATATACGAC AATTAGAGAA
58501 GTTTGAAAAG GTACTATGTT TATTTTAAAA CAGAGTCTGT CTGGGGATGC TGTGTAGGCT
58561 GGTTATTTGA AAATATTTTA ATACCGATGC GTTATGAAGG TGTCTGGTAA AAATTGTGCC
58621 GCAATAAAAT ATGTTAAAAG TCTCCGCGTG GCCCTCTTCT TCTCCAACGT TCTCTGTTAG
58681 ACGAAAGGAT TCCAAATAAA TGGGTTTCAG TAAGCCTTTC ATGTCGTCTG CGTTTTCGTC
58741 CAACAGCTTA TGGAGTTTGA ACCGCACTAG GGGCGGAACG CGCAAAATGC CTTGGGGTGG
58801 GTTTAACTCG TGGTATGCCG CCAAGGTCGA CGCTATCACT TCGATCCCAT ACAATACAAA
58861 CGCGGAATCA TTACACGTAT ACCTCTCGTG TAGGTACCTG GCCCGGTCAA GCGCGGACAG
58921 GATAAATGAT AGGGGGGGCT TATCGTAGTT TTCCAGCATA GAAAGAATGC AGCAGAGTTC
58981 TGTTACGCTC GCGCTCGCCG GCTTAAACAG CTGCACGCGG TCAGAAAAAA CGTGCGGATA
59041 TATGTCCAGT GCGTCGATCA TGTCCTCGTG TTCGACCACG CACGCAAAGA CAAACCCAGT
59101 AAACGTTGCC GACTGAAGCC TGGTGCGCAC GTGCACGCCT AAATCAGGAA CGTTTAGTCG
59161 TCCTCGAACA CACAGCTTAG GCGTATTGGA AGCCATGCAA ATGCTATTTT TTTTATGTAC
59221 TTCCAACATC AGTCGCGGAA AAACTGGATA ATTATAGGGC CCTCTGGACG AGCATTCTGT
59281 TTCAGGCATT ACCGCCTCAG ATCGGTGCTG CTCCCCCATT AGACGCTTCA GTATTTGGTC
59341 CATCGGTTCC AGGCTGGCTA ATTTGTTCGC AATATTTGTG GCCTATGTCA TTTATATACA
59401 CAGACAGCCT GCTGTCCACA TACAGCTGGT CTGCAATATC GTGTAAACTT AAGGAAATGA
59461 GAGAGTGATT TATGAGGTCC ATGTGAGTCA GTTTGGCGAT TATGTACGTA AACGGGCCTA
59521 AGAGAATGGA AACCGTGTCC TGCGAGTAGG CGATGCTCGT TCGCTGCCCC TGATTGTCGA
59581 CCACGGGTGT GAGGTTGTAC GTTTGAAACA TTTCCCCTTC CCACAGGCTA GTTAGCTCTT
59641 TGAGCATCTC CATAAACGGA GGGACGTATT GTGAAAAAAA GCTGTTTACT ACGGAGCTGC
59701 CTTTAGGGAT GTGACATTCG GATAGGTTAA TGGTTGGGTC GCGCAGACCG GATAAAGCCT
59761 GCAACGTATT CTGAGTGAGT ATATTAATAT CGTGTCTGGG CCTGCCTCCC TCTGGTAAAC
59821 TCTGAAGCAA CTGCGTTTCA ATGCATTGAA TCTTCTTTAC GTAGTCGTCT CGTTCTTTCT
59881 CAAGGTTAGA AATTGTATCG AACTGCTCGT TGATCTGGTT CGTCAGGCAT TTGAACACGG
59941 TATTTGAAAC CTGTTTCCGG AGACCCTGGG TCGTACTTTC CGTACCGGTC CCAAAGAGCT
60001 TGTTTTTGTC CACAATGTTT TGAGAGACGT CTGTGACGAA ATCCTCAACA ACGTCTGTTA
60061 TGCCGCTCAC TGTTTTGTTC TCCGCGAGCT TTATCAGCAA GTTTAGGAGC TCCCTCTTGG
60121 GGTCGGTTTT TTGAGACGAT TGCTCTAAGC GAGCCAATAT GTCTTTATAC ATGTCGTTGC

60181 ACGTGCTACC GACCAGCGCT TTTAGAGGCG CGGTATTTAG TAACTGACAT AACTTTGCGT
60241 GCTCCGCGGT CCTGTGACAT GTCATAATCG CGGTAAATAA CCGGTACAGA GGGCTATCAA
60301 AGGCGACGCG TCCATTGTAA ATTATAGGCC GTGAAACAAC TTCAACTGTT ATATCTTTTT
60361 GTCTGTAGGT GCCGAGATAA GCAGCCTTAT CTCCGTTAAA CTCCACAGAC ACCTCGGCGT
60421 AGTCTGGAAT ATAAATAGAC GTCACATCGG TCATTATCGA GGTAATCTCC TGTGCCAGTT
60481 TGTCGACAAT CCGCGATATC GTAACGTCTC CGGAAGCGTA CGTTTTCCTC ATCTGTTGAC
60541 AGTGTGCTTG TATCCTGGCG ATCGTGTCGC CGTCGGGGGT GCGCCTGACA ATAGGCACCA
60601 ACCAAAGGGC GACCACCCAG TCGACGTATC GTTCGTACGA TAAAACAGTC CTGGTGTGTA
60661 GTACCTGATT AATAGTGTTC AGTAACAGTT GATCGACTGT TAGACGAACG GTACGTGCCC
60721 ACGATTCAAA CGTAGACCGT TTGAATTCCG GCGTCAGGTG CTCGACACCG CGCCAACGGG
60781 TCTTCAGGGT GGTCTCGAAG TTGGTCCAGT CCTTTAAGAC GTTGTCGTAT GTGCAGCTCC
60841 CCAATGCCGT CTTGTACAGG TGTATGAACA GCTGTCTTCC AAAGATCCCT GGATTTCGAA
60901 GACTGGAGTG TAGCGTTTGA CCGCGAACGT ACGCGTACTT GCCCTGTAGG ATCTCGAAAA
60961 GCGAGATGGA GAGTTCCGTC GGATGGACCA AACACGTCGA ACCAAATCCG GGGTTCATCT
61021 TGAACATGAC GTCCGATGCC AAAGTCAGGG GTGTCGTGGA TCACGTCAGT CGCCTGTCAA
61081 ATATAACTAC CAGCCCACCG GAAATGGGTT GGTACGACCT GGCCTTCGAT CCGGCTGAAG
61141 ACTCCGGGCC GTTCTTGCCG TTTACCGTTT ATCTAATTAC GGGAACTGCT GGTGCTGGGA
61201 AAAGTACCAG CATATCGGCC CTGTACCAAA ATTTAAACTG CCTGATCACG GGCGCGACCA
61261 CCATAGCCGC ACAGAACCTA TCGCGTCGCC TAAAGACGTT CTGTCCCACG ATCTTCAGCG
61321 CTTTTGGCTT TAAGAGCCGA CACATCAATA TAGCCGTCAG AAAAGCTCAT CAGACCGGAG
61381 CCGTATCCAT AGAGCAAATT CAGCAACAGG AGCTATCGAA GTATTGGCCG GTTATAGTGG
61441 ACATTATGAA AGAGGTTATG GCGAAAAAAC CCAATGGCAT GTACGGGACT ATATCCAACG
61501 CGAATTTTGA AACCCTCTCG AGAATGACCG GACCGTGTTT ATGGACTTCC AATATTATTG
61561 TAATCGACGA GGCCGGAACC CTGTCCTCTT ACATACTCAC CACCGTCGTG TTCTTTTACT
61621 GGTTCCTAAA CAGCTGGCTA AATACCCCTC TTTACCGCCA GGGGGCGGTT CCGTGCATAG
61681 TATGCGTCGG TTCGCCAACG CAGACAAACG CGTTCCAGTC AACTTACAAC CACGGGACGC
61741 AAAAGACGGA GATATCGTCG TGCGAGAACA TCCTAACATT CATGATAGGG AAGAAGGTCG
61801 TATCTGAGTA CGTACACTTG GAGAGAAACT GGGCGCTGTT TATAAACAAC AAGCGCTGCA
61861 CCGATCTGCA GTTCGGACAC CTTCTAAAAA TTTTAGAGTA TAATCTTCCC ATTCCTGACG
61921 AAGTCATGAG TTACGTAGAC AGATTTGTCG TTCCTAAAAG TAAGATTATG GATCCTTTAG
61981 AATACATTGG CTGGACCCGA CTCTTTTTGT CACATAGCGA GGTAAAGGCG TATCTAACAA
62041 ACCTACACAC ATGTCTAACG CTAGGGGGCG ATACCAGGGA CACGAAGCTC TTTACCTGTC
62101 CCGTGGTGTG CGAGGTGTTT GTGAAGCCGT TTGAGGAATA CAAGCGGGCC GTCAACCTCA
62161 CCAACCTCAC CGTGACCGAG TGGGTGACAA AAAACCTCTT TAAGTTAAGT AATTATTCGC
62221 AGTTTGTGGA CCAGGACATG TCCATAGTCG CCACGGAATC CACTGAACGT TCGACCCAGG
62281 TTACCTTTAT CACCAAGTTT GTTAAAAACA GCCACGTGTC TTTAAATGGA AAAACAAAAA
62341 AATGCATATG TGGGTTTCAG GGTACGTACT TTGAGTTCAA AAGAATCCTA GACAGCGAAC
62401 TCTTCGTGGA GACGCATTCG CAAGACCGTC CGGAATACGT ATACGGTTTC TTAAACACGC
62461 TACTGTACAA CGCCATGTAC TCGTTTCACG CGTACGGCGT GACTAGGTCA CACGAGAAAT
62521 ATCTGCAAGA CCTAAAATTT GCACCCCTCC CGGCCGCTCT GGCAACCGGG CGTGTAGACC
62581 TTCAAACGGT TCGTGAAGAG TTAAACCTGG AAGACGACAT CTTCTACCAC GTGTGCAGTC
62641 CCCCGCCACC CGCGGGTATC ACCTCCCTCC AGGTTTTGGT CGACACGTAC TGCGCCCTAA
62701 AGGACGTGTT CGCCTCCAGA ATAAAGGTGG CGTGTCGCTG GTTTGGCGGG GAGTTTGAGA
62761 AGGAAACGTT TTCCGCGTTT ACGGTTAACA TGGTCGTAAG GGACGGAGTT GACTTTGTCT
62821 CCCCTTCAGA ACGTCTCAAC GGGCTGTTGG CGTTTGCATC GACCGTTGAA TCGTATAAAA
62881 TTAAGGGGTA CACGTTTTTA CCGGTAGCGT TCGGTCGCTG TCAGGGTTTG CCACTCAGTG
62941 ATGACCTCAG GAAGAAGATG CCCTCCCTGG TCGTGCAGGA CTCTAGCGGT TTTATCGCGT
63001 GCCTAGAGAA TAACATAACC AAATTGACCG AAACCATGGA GGACGGGAGC GTTTTCCAAG
63061 TGTGCTGTGC GGGGGACTAT GGGGTCAGCT CAAATTTAGC CATGACCATC GTAAAGGCAC
63121 AGGGAATGTC GTTGGAGCGA GTAGCCGTAG TATTTGGGTC CCACAAGAAC GTCCAGACAA
63181 GCCACGTGTA TGTAGCAATA TCAAGGGCTG TTAACTCAAA TTATTTGGTC ATGGACAGCA
63241 ACCCCCTTAA AACCCTCCTC AGAGAACCAG TCGATAACAC CTCCGCCAAG CATATAGTCC
63301 GCGCCCTCCA CAACCCAAAC ACAACCCTCA TCTACTAAAT AAATAAAAAC AAGCAAATGG
63361 AAAACATTGT GTTTTTATTC AGTCCAACCA CGGCCACGGA TAGTTGTCAT TTCCACACAC
63421 CGGGGGGGTG GCCGCCATAG TTTGACGACC AGAAACGCCC GGGCCCTCTG GGTTAGAAGT
63481 AGCCCCGCCC CCCGTGGTCG AGGTCATACG TCTCTTAGGG CCCTGGGTTG TACCAACATA
63541 AATCACGTCG TCATCACCGT TGCAGTCCCT GCGTGACGCG CGCCTCTTGC AGGCGGCACG

63601 AACACCGCCG CTTGTAACGT GATCTGCATC GTCGTCGCTG TCTGAAGATG AACACAGGTC
63661 AATCACATCA GTGGCGCCCG TGCCCTCATC GGCGTCTGAT TCGTCCCATG TCTCAACAGA
63721 GCCTCGGTCG TCCACTTCGT CGTCATCTGA AATAATAATG GTCTCTTGTC CTGCAGTTGA
63781 TCCAGCAACA ACAGGTTCTT CTTCATCGGA AGACCATCCG CCCTCTGCCG GTGATCTAGA
63841 TATATCACCT GACCAGGCGG GCGAGGCCCG TGGCGAAAAG CCGTCAGCTT CTCCGCCACT
63901 TTCATCAGTG TCCTCGTAGT CATCTATAAA AAAATCGTCT TCGCTGTCTG ATTCAGATGA
63961 CTGGGGTGCG CGGTGTTGGC CAGTAACCAC CGCAACGGGC CGCGGACCAC CAAACGCGGT
64021 CAGTGGCGCG GGCGCGCTCG GTGGTGTGGT GGCACCTGTC GAAGTAATGA CGGTAGGTAG
64081 CCGGCCGCTC TGGTTAAAGG ACTCACCGTC CCGTGCGTGC GGATCTATTC TAGGTGGGGA
64141 GGAGACCGGC GGTTCGTCTT CTTCATCCAC ATCACTGTAG TCGCTGTTCT CCGAATGAGA
64201 GTCGGTTTCG TCTGGCGTTT TTGGAGTAGT GGGAGGGGAA TCTTGGTCAA AAACATCCTC
64261 AGGTGCAGTC ATGTCAATCA CTGGGGGGCC TCCGTACCCG GGCGGTACCG GCAACCCCGT
64321 ACGGGACCCG GCGAACAAAA AACGTCTTGT TCTCTTTTTC CTAGGTGCCC CGGGAATCGG
64381 CAGCATCCTG GGAGTGGCCG GCGGGGTCCT TGACGGGTCG GATAAGAACA TAGCCATGGC
64441 CGAAACGTTC ACCTGTAAAA CGCACTGCCT ATCCCCGATC AGTTAATATT CCAGTCAACG
64501 GCACCCCTCC GATTCTGAAC TAGATAGTCA TTCGCAAGTT TAAAATGGTT GCAGCCTAGA
64561 AACTGCGGCC AGGGAGAACC GGTTGCCGCG TGGGCGGTAG CCAGGGGAGA CGGATGGCGC
64621 GCCTTGAGGA CGAGATGTCT CTGTCCGTTT ATCAGCGTCG CCTTGTCTAT AGCCTTTCTT
64681 CCCCACAGCA TAAAAACGCA CCTCTGAAGC TTGGCAGAAA GGCAGCTTAT TATGTAGCTG
64741 GTGAACCAAT CCCAGCCAAG GTTGGCGTGT GACCCCGCCT TCCCTCTTTC CACCGTCAGA
64801 ATGGTGTTTA AAAGCAGAAC TCCCCGTTTT GCCCAACAGT CTAAGCACCC GTGAGAAGGA
64861 GCGGTGAACC CAGGTACAGT ATTGGCTATC TCTTTAAAAA TATTTTTGAG GCTTGGAGGT
64921 ATACTGTAGT CCGGAGCAAC GCTAAAGGCT AGTCCGGTGG CTTGACCGCG ATGGTAAGGA
64981 TCCTGGCCAA GGATCACGAC TTTAATATCC TCCGGTTCGC AACAATACGA CCACCACATA
65041 ATTCTATCTA TCGGTGGATA AATCACGGTT ACGTTGCTCA TATCCATAAC GCGCTTCAGC
65101 AGAGCCGCGA GCTTCTGTTT TAAAAATGGA GAAAGATTTA AAAATTCAAG CCATGAGTCG
65161 CTAAGTAGCA ACGTTTGTGT GGAGGGTTCT TCTAACACTT CTGGTGACAT TTTACTCCAA
65221 ACTATTGTTT TAAGCCAACC CTCCATAAAT CCCCGCGTCT GGTAATTACT TTACGTATAC
65281 GTGAAACTTT TGTACTTCTG TCTGAGAAAT CCAACAGAGG TGGCCGCGTA TAAGCAGTTC
65341 CGCTTTCACT ACCGTTTATG TGAAAGTCTT CAAACTCGGC TAAAAACTCA TCAAGACTAT
65401 ATTGCATATA ATTCAACAAT TGAAGCTCTT CTCTGGGGAC ACTGTAATTT AACTTGTACA
65461 AAATCCTTAT AAAAAACGCC CTAAGGTGAA ATCCATTTAC ACATATTTCT GTGGTAGAGC
65521 GATCAGCTTT ATATCGCAAT GTTGCTGTAC ACAGCCCTGG ACATAAAGCT AAATTTGTTA
65581 AATGAAAAGA AGCAATCTCA AACGGCTCCG GAAATATAAA GTCACTAAGT GGTATTATTT
65641 TACAACATGG ATTTACTACA AGCCCGGTAC AACATGTAAT TAAAAAAAAT CCACATGTAA
65701 TAAATAAACT TAAGGTATAC ATACTTCTCA TGTTTTATAC ATAATATGAA CGCTAATGCT
65761 GTACTTATAT AATATGTATA GAATTATGCC AATACAGTCA TGGTAAGTAT ATTTTAAGGC
65821 ATAATGCAAA CATCAAATAA TATAACACAA AATGCACGCT TCCGGATATC CGCCCAGCTG
65881 TTAACCCGAA AATACAGAAA TGACTACACA AACACACCTG AAACCAATTT TATTCTCAAC
65941 ATATGCAACA ATATTTCAGG GTAACACCAT GTTAATAAAA TACGCAGCAT GCACATTTTT
66001 AGCTAAGCTC CTCAAAGACA ATTTCCTCTT CAATTGATGA CTCATCAACC TCGGTTTCCG
66061 TGTTAATGTC GGAAATATAT GAGTCCAGGA TACCCTCGTC ATCGCCACAA ATTAATTCCA
66121 GGGCTTGTAA AATATCATCC AGTGAATCAG CTGCTAATGT CAAGGATGTT GTTTCTAATT
66181 TTAGCCCATC CAAATGGTGA GTGGGTGACA CTTTGGGACA AGACCCCGTA AAATTGGCAC
66241 GCTCCACACA AACGGAAGGA TGCATTTGAT CGTTCACGAG GGGCAGCATA TTTATTTTTA
66301 TAACATCAGG AACCGGACCC GTCAATGACG CACATATGTC CATTAAGATT GGAGTCGTGA
66361 CACTGGTTGA ATTGACGGTG AAAAAAAATG TATTCGTGTC ACATGCGTAA CTATTTTTTA
66421 ACCACACGAG GGCGGAATAA AACGGGTCCC CGGCGCTAAG GCCGTGAATG CGGCAGGCAC
66481 GGCTCATTGC CATTTCCAGA TTTGTTAAAT CTGTTCGAAC ATACGAGTCT AAAATCAAGT
66541 ATAACCACGA CAAGTTCAAA CAGGAGGCAA CGCGGCTGGT TTCCACCCGG ACATCCCCGC
66601 TCACAAACCC GAACTCTCTA GGAACGCACA TCAAAGCGTT CAGACAAAAT TCCGAAACCG
66661 AAACCCTAGT TCTGAGCACC ATCTGGTTAT AAACTGCGGT TAAGAGACGA GCGGCCACAC
66721 ACTGCTGCTT GTGAGGATTC AGCTTAGGTG GCCTGCAGGC TTGCTGACAG GCCCGTAGCT
66781 GCCTGGCGGC GCTCGCGCAC CTTTGACAGC CGCACGCCAG TTCCAGAGCG AGGCAGAAGC
66841 GCTCTTTGCA GTCGCGCCAT ATCTCGGCCA TAGGAGATTC GCTCGCGTGA TCTTTTAGGT
66901 GCATATATAT CTGTTGGGTA ACAAACCTCA CTCCCTGTAA AAGGGGAATA AGGTCCGTTC
66961 TTTCTATCGT TTTCTCTGAA TCTATAGTTA CATCTTTAAG TATAACCAAA GACGCTAAAA

67021 ACCCAGGATG GTCGACACCT TTAAAAAATC GCAAAAGCGA CCGGATGGCT CTATCCGCGT
67081 TTCCATGTCG TTCAAAGGTA GTCACGATGG ATCTCCAGTT AGACTCTGTC TCCCTGTTAA
67141 CACCCTTAAC TGGAATGGAA ACAGCCATGA CACCGTGAAC TTCCTGATGT CTCCTAAAAC
67201 TAACCCCGAA ACAGAGCTAA ATACCAATGA CTGTCACCCC TACCCCAAGC CACGCCCCCG
67261 TACTATTAGA CCAGGGTGAG TAACCACGCT ATCCTTTAAA AACCCATACG TGGAGTTTGT
67321 AAGGTAAACA GCTCGTGTAT ATTTCAGACG CCTGTCATAA AATGGATACC GACGACAATC
67381 AGGTAATTAA ACTTTTTTTA TTCAAGATTC AGGAGGGCGT GTTTACAGTG GTGTAGGTGG
67441 GAGCATATCT CGCCAATGGG AATGGCTGAT GAATCCACAC TTAGTGCTCG GATCAGGTGC
67501 TCTGTGTACT TTACTAGTGC CGATGAACCT AAATTTGGCG TGTTTAGGTC GTACCAACCA
67561 AAAAACTGCC TGGCCTCGCC TTTTAGAAGC TCTATGCAAG CTTTAATCGT GTCCAATAGC
67621 TTGTCTTTAA AAATGCAACT CTGGTACATC TTTATGACAG TGGTCCAAAA AAAACAAAGA
67681 TTTAAAAACA CGTTAAACTG CGTGTCCTGG TAGTCCTCGT ATATAACCTC CTCAACAAGA
67741 AAAAATTTTT TAACCAAACT CGCCAGGTAC TGAAACGATG GTACGGACAG GTCGTGAAAA
67801 GTGTCTATCA TCTCCTTCAC CTCCCCTTTG AAGGTTTTGG TTACACCGAC CATGTGAGAT
67861 AGGCACCAGC TCAGGGGGGA GGTCGGATCG TGAGGGGGTA ACAATTCGTG GGCGTGGGGG
67921 TACTCCAGTT TTAGGATCTC CGGCAGGACG CGTACGAGTT CTTCGTCAAA TCCTACGCAG
67981 CCCGCATCCA CCAGGGGCAG TACTGAGTTA ATAAGTCCCC GAATTCTATC GTTCGTAATT
68041 TTGTGGAGCT CCTTTAGACA GTAAACCATG TTGCTCCGAC ACTGCGGTAC GATGAACCGC
68101 TCTCCTTGTC TGTGTTGCAT GAGCACCTGG AGTATCTTAG ACAGGTATAG GGCCCGTTCC
68161 ACTCTAAGGG TTGCCGCTTG GTAGAGCGGA TTGCGGGCAT TGACGGTCAG TGTGTTCATT
68221 TCACCGACTA TCATGCCGAG TGGCGGATAG TGGAACTTGT ACATATGATT TAAAAGATGG
68281 TCTTTCCCGT AATGGCGACT CATGGTCGTG TTTTAACGTT CACAGGGTGT TAAAAAACTC
68341 AGAAGGTCCT GCGTGGACAA CTTTTGCGGA CTCTCCGAAC CGCTCAGGGA TAAGCTGTAC
68401 CACGCGGTGC TTCTTATAAA CGAAGCATAC GTCAAAAAAC ATGACCCACG TGACCTAAAT
68461 AAGCTCACAG AAGAAGTCTG TCGATGTATT GTAATGGAAT GTGCCTCCCT GGGTCCCATC
68521 TCCGGGCTCA TCGCGGATCT GAATCTCTTT AACCTGTTCT GCCTCTACCG GGGTTCCCGG
68581 GTTAAAACCC GCGGGGCCGC CACCTGTAAC GTCCCGTGCG CAGAATGCGC CCAGGGCATC
68641 GTGAGAATTC TCACCGAGAG GGCCCTTTGT TGTACCGAAA AGATGTTCAT AGCGTCTGCC
68701 TGCAGCGGGG TCGTTATACC GCCCCAGTTG GCCAGAGTTC TGCACGACGT CTACGCGGAA
68761 ATGAAGGCCA AGTGTTTGGG GGCGTGGCGT CGTCTCATAT GTTGCAGGCG GCCTATCATG
68821 GCTATTGCCG ACTCGGTCCT CGTCACTTAT AATACCCTGG ATGCCGAGGG AAAACTGGAA
68881 CTTAGACTCA AAGCACTGTG CAAACTTGTT TTTCAACCCA TCTTTCTTCA AAGAATCTTA
68941 GCCCCTATGC AGCTACTGGC CAACGGGAAG ATGGTTCCTG ACAATTATTT TACCATCACC
69001 GGTACGGCCG AGAAGAGGCG CCCTGTCGTG ACTGGAAGTA CTAGCGGGAT GACGTGTCCG
69061 GGAAGCAGCC TTGTCCCCGA TTCCTTAATC CTGCCAGTAT GCGAACCGGG GCTTCTCCCG
69121 GCACCCCTGG TTGACCTCAG TAATGTCTTA GAAAATCCAG AAATCATCCT CAGCGCCCCA
69181 CCCCTGAGTC AATTTGTCAT CACAAACACG CACCCCAGTC TGCCTCAGTC AGTCAGCATT
69241 ATTACGCCAA CCCAGGGCGT TGTTCCCGGC CAATGTTTTA TGGACACGTG GAAAGCGGTG
69301 TCACAGAGCA TTCACCACCA GGCACAGACG CCTATTTTGG CCGCCGCACT AACCGGTTCG
69361 ACATCTGCGG CCCCTGGCCC GCATATCGCA TGTTCCCCAG TTGCCGGCAC GTCTCGGCAG
69421 GTGGAAGGGT CCGCGGGCGT CGATTGCGGG AAACCAGCAT GCGTTCCGCA GCCCGCGTTA
69481 CCGCCCAATG TCCCCGCCAA GAGGATGGAA ACGGTAGCAC AGTTGGGAAA CGCTCCCGTA
69541 AAAAACGTCC ACATCGGAGG CCGCGTATAC GCTCCACTGG TTAATATACC AATAATAGAC
69601 TTAACGTCCC CGTCAGGGTC CGGCCAGAGT CCGGCCGATA TCGCCAACAC TCCAGAGTCC
69661 CGCATGGCGG CCGGCTCTCC GCCCTTCGCC GAAACCGCCG CAACGGTCCC CGCTAAGAGA
69721 AAGCAGCCAC GCGAGGACGT GGCAGACAAA AGACTGAAGG GCGACGTTCG GGGCGCCGCA
69781 ACAGTAAACC ACCCTTTCCC GGGACCGTCC GGGATGCGCG TTCGCGAGCA GGGCTTATTC
69841 GATTTAATCG AAAGCTCCAC GGATGTAACC GCGAACGCAT CTGGACCCAA AAACGACGAC
69901 GACATGCTAG CGGCTATCCT ACAGGACCTG TATGGACTAC AGTCCCCCCC GGCCATCGAT
69961 TCCCCCTCCA GCAACTCGGA CAATGAGGAG ATATTTCCAG AGGTTAGTCC GCCATCTAGC
70021 GGCCACGGAT CGCCTTGAAA GATGACCCCA AGACGCAAGA TCACACCACC GAGACAGTCC
70081 CACGCAACCC GGCGGATGTC TAACCGTAAA AATATCACCG GTTGGAAAGT CCCTTGGTGC
70141 GACAAAAATG AGCTTCAAAC TTACCCATTT TCTGTGCTAT TTCGGGACAC TAAAAAATTG
70201 TCGCCAGTGG CGACAATTTT TAAGTCTGTG TAAATGTCAC TCTTTCTCAA GTTTCTGTCC
70261 TAGAATTTTC ACATCCAGTT TCATATCATA TAAAGACACC TGTTTTGTTT TCGTGAATCT
70321 GGCGTGTTTT TATCTTGAGA GGCCCCCTAA AAACATGCCT CGTGTGAAAA CACAACCCAA
70381 GAGACCCCAA GTGCTTGAAT TTATGCCATT AGATCTCCAC GGTGGAACAC ACACGGAGAT

70441 GGATTCTCAA AACCTGTGTC CTGACGGCCA GGATCTGCTC GGGTCTTATA TCTATACGGA
70501 GAATAACGGC CCGTTTTCCC AAATAATGCA CAATGGACAG AGCAATACCG GGACAGGTGA
70561 AAGCTTCGGC AGCTACGCTG CCGGCGACGG TTTTCTGGGC GGTTCTGTGT CAGGGATGTA
70621 TGGAAACAAC ACCGGAGAGG GCGCGTGTTC TAAAAGACCG TCCGCGTGCC GTAAACGCTC
70681 GGCTGCACTA ATTCACGCGG CGTCCGAGGC GTCTGTGGCC GAGCAAGGCA CCTCACAGGG
70741 GGCACATGCC GTATCTGACC GGATAGGCAG AGACGGTGGC GCTGACAATA GACTACTCAA
70801 GGTGAGTGCG CGGCTGTCGG ACAAAACAAA GAGCGCCCTT CGCAGCCATC CTTGCTTGCG
70861 TTGCTATTCT TTGATGTTTA ACACGTAATG TCTGTTTATA TTAACCCTGC CATAGCCCGA
70921 ACCGAACACT GTGCGATGCG ATTATAAAAC GCAACACTCG GACACGTCAG ATGGACGCGC
70981 AAAACGCGCG CGGCAGTGCA AGACCTCAGG AAAAAATAAA GCCCAACAGC TGCAACAGGT
71041 ACAATAAGCA AATTTAGTAA TTTGTATTTT TTAAGTTTCC GAGGTTTTCC CGCTGGCCTG
71101 CGGCACCGTG TCCGTGCTTT AGTTGGAATA CCTCTCCTAA CAAGCACGTA TTCCAACAGG
71161 AGCTGGAAAG TAAAACAACG CAACTGACCG TCCTCACCGC TCAACACGAA GCCCTTCAAA
71221 AGCACGCATC ATTTTTACAG AAGATGATGA TATTGATGTG CAAAAACGGG AACAGTAAAA
71281 AAAGCACGTC ATGATTTAAA GCCACCTCGG GTTACTAGTT TAAAACTGCC TAATTGATTC
71341 TATTCACCCT TCAATAAAAT AAAATTTTAA TACGTTTGCG GTTTGTCCAT TTCCTGTCTT
71401 AAAACTATTA AACAATTCAT TTGGGAATAA GGATTTATGG GATTAAGGGA TTAAATTTTT
71461 TTGGATCATG GGATTTGGGA ACATACGTCT GGGATGGAGG TTATGCTTCA TGGTCTGGGT
71521 GGCGTGGATT GCACGGGGAC GGTCGGTGTG CCCAACCTGG CACCTGACAG ATGGGAAATA
71581 CGAGGCGGTA TACAGGCACT ACCTCGAAGA GTGCCGCAAA CATGAAGGCT CGGGGAGCCT
71641 GGACGGTTCC GGACAGACAA AGGGGTCTGG AACCAAAGCA ACCACCGAAG CTAATATATC
71701 GATAAGACCT AACGTTGTCA CATCAGGTCA AAATAAAGAG CCGCCTGGGA CAGCACCGAG
71761 GGCCGAATCA TCACACGACC TGCCACGCAT CAAGCAGGTT AACGCTCTCC GATTATCAAC
71821 CCCGGAATTG GCGCAACCAC TCCCGGTAGT AAAATCGACT CCGCGCGAGT CACAGTCAGG
71881 TGGGACACCC TGGAACGCGC GCCCCCACGC GTTCATTATG CACACAAACG ACATGCTCAA
71941 CCCATCTGTG GTCCTGTCTT TCAGAGCCAT CCGTGCGCGG TCCACACGCG ATACCGAGCA
72001 GTCCGTTCGC GATCGGAACA CGGTCACGAC CAGCTATCGT ACCCCTGGCC GCCCTTCCCT
72061 CTTTCAAGCC AGACCCTCGT CTCACGGTGC GCGTCTACCC CCTTCGCCCC GAACGATGGC
72121 AAGATACGCC GAGTCGCGAA CAATATGCGA CCAAAATTGA CCGCAAAGAA TACCACGCCG
72181 TCCGTTCGAC GCAGTGCCCG TATGGACGCT GCACGGGACC TCAAGCGTCC GTGTCAGACC
72241 CATTCCTCCC TTCGCCGACG CCGTGGCACC TGGTTCACCC CCTGTGGGCT ATGGCCGGAA
72301 TATTGGGCCT CACGATCGCC TTGTACACGA TATATCAAAT ATATGCACAT AGAAACTACA
72361 CGACGCTTCA CTGCGAATAA AGGTTTATTT ATTTTGCACA CTAGTCCGCG TCGTTATTTC
72421 TGGTCCTGGA CTGGGCGCGC CTCCTCCTGA GCGCCCCGTC CGCGCTGGCA GCGATTCCGC
72481 CTCTGGTGGT CTCGTCCATT GACACGTCCA CGCGGATACT AATTCCGGCG AGGGCGTCTT
72541 CCATCTGCCC GCGCGTCACG GCCTTGGCCG TGGCCGCCCT CACCTTAGCT TCAATCTTAC
72601 GGGTTGCCAC GTTACATAAG GCCGAAACGG CAGCTGCGAT TTTGGCCTCC TTTTCGGTGT
72661 TAGAGATGAC GGGGTCGGAG CCGGACCTGG TGGGATCCCC AGAATTTATC AGCTTTTTGA
72721 GCTGCTTGTT TTCAACGGAC AATTTTTGCA CCTGAGCGGC TAACTCCTCC ATTGTCAGCT
72781 CTTTTTTGGG GGCGCGCGTC TTGGGACGCG TGGAAGACAT GGCAGCTACT ACGCATTACA
72841 GAGTACAGCG GGACGTTTAA ATACTTTCAG GCGGCCCCTG AGGTCACTCG TGCGCCACGG
72901 TTTCAACAAA TTTAAAAAAG CACATATATG TTAGGTACAG AAATGTTGCA ACCGCGACCA
72961 GAACAGAGTT GATCACCGCC CATATACTCG AAAAACTAGA CAGCACGGGC TGGTACGTGT
73021 CCGCCGAGCA ATCATAGTCA TAAAAAGAGG GCGTTACGGA GGGAGCCGCG GAGGGGTTTG
73081 TAGTCACAGG AGGCGTGCCC GTTCCTGTAG TCGAGGAGGT GGTGACGGTG GCGCACACCG
73141 GCAGACACAG ATATAAGCAC GCCAACAGGG CTAACGCGAG CACAATTGAC CCTGTCATTT
73201 TCAAGCCGGC GTTCGGCAGG CAACCCTCCA CTTTGGCACG CGCGCTCTCC TTATATACTA
73261 CCAGTTAAAC ATTATGGCAG AAGTGACCGC CCATACGGTT CCGTACGCCT TTGACTCCTG
73321 TAAGTTTGAA ATAATACCAA AAAACAACTC ATCTCGTATA GCGTTACGCA ACAAATTTCC
73381 CGTCGTGGTC AAACCGGGAG AACCTCTAGT CGTGCCCCTG GGACTAAAAA TTATCCGCGC
73441 GCCCCAGTGC GCATTCTTTC TCAGCGGAGC GCCGACGGAT GAGGTGTATT ACCACACCGG
73501 ACTTATAGAT CAAGGGTATC GTGGAGAAAT AAAACTCATC GTTCTCAATA AGACTAAACA
73561 GGTCGTCACA CTTTACCGGG GAGAAGTCAA CGTCTCACTC ATCGCGTTTA TGTACGCCTC
73621 GCCCGGGCCC CTAAAATGCC CGATACTGAA CCTACCACAT TACTCTCTGG ACGCTGGATT
73681 TGATGTTACG TCGCCACATG CCATGACCAT ACCCCCAACC GACAGAACCC CGTTTACGTT
73741 GAGTTTGTAT TATAAGAGTC CACAGCTCAG CACCCCCCAC GTCCCACTAA TCGTGGGGCG
73801 CTCTGGACTG GCGACAAAGG GGCTTACCGT CGACGCGACC AAATGGACCC AGTCGTTGGT

73861 ACATCTGAGG TTTTATAACT TTACCAAAGA GCCAATAGAT ATACCAGCAA ACAGCCGCAT
73921 ATGTCAGGTA GTGTTTATCC ACGAAGACCA CGTCCCAAGC GGGTGGAACA TTCTTAGATC
73981 CCGCGTACAA CTCGGCAGCA CCCTCCAGAT ATCGTGGGCC AAAATCAGGT TTACCGACGT
74041 GGCCACGCTC CCCAAAACCC ACCCGCTCAA CTCCCGCCAC ACTCAAAGCC AAACCGAACC
74101 GGAGACCGCC CGCGGCGCAA AGGGGTTGGG GTCGTCAGGG TTATAATTTT AAAACATGGC
74161 ATTATTTTAA AGTTGTTTAT TTTTAATAAA ACAATCCAAA TTTACGTTAA CCATCAGCTT
74221 GCGCATCCTG TATTATTTTG TCTGGCACGA TATTATCCGC CACCGCCAAC TGGGGCGTCG
74281 GTTGTCGAAT CACCGGATGG GAGGGGTTGG AGGACGCTTC GGTGTCTACC AGCACATTAT
74341 TAATTTCCGC AGCCCACGTG ACAGAATCAT CGGACGCGAT CGTCGTAGCC AGTTTTTCCA
74401 TTTCCAGGAG TGACACGGTC TCGTCGGCCC GTTTATTCAA AAGCGCGTGA AAGGCGCTCT
74461 CGTCCACCGT AATATCCCCT ATAGCCAGGG ACAACATGGC GTTCCAGATA ACGTTCTGTA
74521 TAGACTTTAG CTTAGATATT TGAATCCTAA CCGTCTCCTG GCACCCCGCG GGTACCCTAG
74581 CCGCGTCAAA ACGACGAAGA TACTCGGTTA TTTTATTGGA TTGAACGGCC AACAGATACG
74641 TCTGCCGAAG GGCGGCGGTG CACATCTCTG CCTGCTTCAC ATCTCCCATA TCCACTCCCG
74701 GGGGCAGACC CATGGCTATT TCGGTCCGCC AGTGAGTAGG ACACTCGTTA AGAGGATCTG
74761 GCAGGGGACG ATATTGACAT CCCACGCGTC CAAACGGACT AAACCCGCAA CACGTCCAAA
74821 ACGAACCAGG CGCGGCCATT GCCAAACAAA ATGGTAGATG AAATTAGGGC AATTTTCTCT
74881 ACTAGTGGAG ATATGGCCGA AGTAATTACG GATATACTGA CTGAAACGCA AGCAACGGCG
74941 TCCTTCTTCT GCGTGCTCCA CGATCGGGGC GACGCGCCTA TAAATACTCC ACATGCCGTA
75001 ATTAAACTCT GCCTGCCCGC CAAGCGCCCA GGCGGCGGGC CAAGGTGTTT ACCGTTGATG
75061 GTGCTGAACC TACCGGCGTG GCAGGTTAAT CTATTCCTAA CAGGTGACGC ACCATTGACC
75121 TCGGATAACA TTAAAGACCG CATTGACCTG GCTCAGACCG AGGAAATACT CGAACCCATA
75181 TTAAGCGTAC TGGCATGCAA ACGGTCCGCG CAGCAGACCA AACATGACTC GTTTAAATCT
75241 AAGGTGGCCT GGTTCAGGGC AAAGTTCGTC TCGGCCCTAA GAAAAGTGTA CAAAATGACC
75301 CCATCCCCCT ATTGGATGAT AACGCTGCTG GGCTCGTTCG AGGCATCGTT TGTACTGGCC
75361 GGCACGTTTT ATTTTTTCCA GTCTCATATA TGCACGGCGG AGACGCTGGT GCATCTAACG
75421 AGGTTATTCA GCTCCAGTCA GGGCCAGAGC TTGGTCACCG TAAACACCTA TGACGAATTA
75481 GGGCGCGTAT TCGGGCGGTC AGATTTCCTT GGAATTGTTC CAAATTTTTG GGCGTATCTA
75541 AAATACAAAA TGCAACAGGA CGACGTGGAG TCCAGGGCCA TCGATCAAAC AATCAACTCC
75601 ATACGTGGGG GGTTGATGCT GTCACCTCAG GACCTCGTAC ACTTCATCTA CCTGTCATTT
75661 TACGAGTGCA TGAACGCACA GACGTTCCTG TCGTACTCTC GTACCACGTC CAGTTTACCG
75721 ACCCCCGCAA CGGTTAACCC TCCGCAGCTG TGTCGGCGCT TAGAAGCGGA TTTTAAGGAG
75781 CACGTGATGG CATATTACAA TAAAGCAAGC TACCTGAGTA CTTACATAAC CATTTTAACC
75841 GTACCCGCTC CTCTCCCGGA CGGATACGAA AACTTTCAGG AACTAGCATG TCAATACTGG
75901 TGTGGACAAT CGAGAGACGT GGCGGAAATC ATGACTAGAA TTAATGACCA GTATCCACAG
75961 CTGAATCTAA CAAAAGATTT ATCCGGTCTC CTTGACCTAG CGGCCCTAGA TCAGTATTCC
76021 GGAGGCCCCA AGGAAAACCT TTTTACGGTG GCTTCGAGAA TTCCCACATA CAGGTGCGAG
76081 TTTTTGAATA AACAGTATTT TGTTCTAATG CACGCAGACT GTATAGACGC GTACTGGAAA
76141 CAAAACATTA TCGTGCCCGA AGACGCCCAA TTGCAAGGCC TGACGGATCA AGATCTCACC
76201 TCCAGGATAT TTTACTGCGA CCTCGGCCTA TCTCTACCTA CATTTAAACA ACAAATTCTA
76261 GTTTCACGCC ACGAGTATTT CAACCCACGA CTCCCAGTTT ATAGATGGGT GTTGGATTTT
76321 GATCTAAAGG TAACCGAAGG TAGACGGACT CTAAACGATA TCTACAACAT ATGCGTAACG
76381 CTGCGGCAGG TCATATTAGA AACGTTGCAG TTAATAGGTC CACTAAAACC GAACCACCCC
76441 GTGTATTTTT TTAAGTCGGC CTGTCCAGCC GTAACCTGGC CGGATGATAT CTCAGACACC
76501 GCGTTCTGTC ACTGTGACGC AAAAATAGGG ATGAGAATTG TAACCCCGTT TCCGAGCGGA
76561 TACTGTTTGG TGGGGTCCGC TCCACTCGTG TCCCTGACCG ACATTTTAAA CCGCGTGGTT
76621 AAACTGGACA CGCGGCTAGC CTCAGAGTAT CCTGGGATCT TAGAAGATAA AGGACCATTT
76681 GACTCTGGCA TTTACGCCAA GGGGCGATGC GTGCGCGTAC CTCACTGCTA TAAGGTTGGC
76741 CCGGGGGGCG AGCTGTCACG GCTCCTAAAA ATTATTATCT GTCACCCCGA AGAGTCAGAT
76801 AAATCGGCGT ATTTGAAAAA CGCGTTTAAG GTTTCTAATC TGCTACATCA CGCCCCTGGA
76861 GACTCTGTCA CCAAAAACGG CCACCTGGTG TACGCGATCA CTGACGAAAA CGAGGGCTTT
76921 CTAGAAAGCA AAACCAAGAA TAATCTTCCC AAAACAATCA CGGACCTGGC CGAAAAAATT
76981 GAGCGAACCA CAGAAAAACC ACTAATTGAT TGGGCAGCGA CCGCAGTGTG GCCAAAACTA
77041 CACGATACCA TACAGCGTTT CTTTCCGGAT GACCGCATCG GCCAATTTGC ATCTGTGAGC
77101 TTCATGCACT CCGGAGACAA CATTATACAA GTTAAACCAC AAAAAGGAAA CAACTTTTTC
77161 TGTATTAATC ATAAGCACCG AAACCACACT CAGACAGTCC GTGTATTTTT AACCCTACAT
77221 TCCACAAAAG AAAGCGAGGT CACGGTGACC TTTATGAGTC AGTGCTTTGC TGCAAAATGT

77281 AATCATAATA GCCCAACTGC ACATTTTTCA TTTATGGTAC CCATAACCGG CACGTAAGAG
77341 ACACGAATAA AATCATCCGT GATGGGCCAG CTCTCCGGCT ACTTTGTTAA TTTTAACACC
77401 GCGTAACCCC AAAAACAACA TCTGGACGAG TTTTGTTCGC GAACGTGAAC ATGGCACAGG
77461 CAATGGTGAG CCTGGAATAC ATGAAGGATA TTTTGGACGG TAAGAAGACC CCTTGTGGCT
77521 CTTACGACTC CCATTTAAAA CCGCAACTTA TCAAACGGCT CTTTTTATAC GATTTACATG
77581 CGGTATGTGT TTCACGCACT AATTTGTTTT ATAGGGGGCA TATCATCGTC GGACTTCGAC
77641 GACTCGTCCT CGGACGAAAT GGACGACCTG TCCCCAACGC CGGAGCCAGA ACCGTCCACA
77701 ACGCCGAACA GCTTTCCGGA GGGGCCCAAA TCACAAGTGG TGGCGTTACC TAAAATCCGT
77761 AAAGATCTC GATCTGAGAC GCCGGTAAAA ATTGAGCACA GATCTCCACT TAACCGCTCG
77821 CGATCTCGAT CCAGAACGCG GTCCGGGTCC GGTCAACGAT CAAACCAGTC AGGCAGATAC
77881 GTCAAGAGAT TCAAACCAAC GGTTGATGCA CCCCGTCATC GAGAACCGTG GCACAGGGGC
77941 GGTAAGGGAA AGGCCCCGTT TATCCGCAGA GACGCAATGG CTGGCCGCGG TCGACGCACA
78001 TACGGCCACG ACTATCGCGG AAAAGCCGCT TTAACGCGGA GCATTAAAGA GTCTATTAAA
78061 AAGATGCACC TTCCATCCAC CATGCTCTCT CGTGCGCACG ATAAAAAGGT ATTCGAGGGA
78121 CTGTTGCCAC GACACCTGGG ACAGTGCTTT CAGGTGTGCC TGCCGGCGCC GCCACCGCTG
78181 CAACCCGAGG TGTTCACGGA TCGACAGCTT ACCGCTATAG TCAAGTCTGG CGGGCGCAGA
78241 GACGCGCTGG TGGCCAAAAA AGTTAGCCTG GCTAAACTAA CAAGCCTATA CAAACCCCTG
78301 CTCACGTTTG TGACGGGGAG AAACAACCAG GCCCACTGGT TGGCGACGCG CAAAAACACG
78361 CTAGCGTCTG CGGGACTCGA GGCTCTCGCG GCCTTTATCG AGGAAGGTCT GGCGTGGGCC
78421 CAGGTGTGCG TATCTCAAAA CAGATCGTTG AACGACAGCA ACCTAGATAT CATTCTAGAC
78481 AGTAGTCAGA GCGTGTGCAC CTGGTTCATC TCCAAAATCA GGCATTTGCA CATACAGTGC
78541 TTCTTGGAGA ATCAGGGTGA GGTTAGCCTG GTGAAGCAGC TAACTTACCT CGTGTGTATA
78601 AATAACCGCC TGGCGGAGGC GGCGAACCTG GCGGGAGAGG TGAAACTGAA TTTTAAACTC
78661 GGAATGCTGA TTGGGTTCGC CCTAACCTTG CCGGCGCTCC TTGCAGAGCA TAAGTTGTCT
78721 GGCGAGAGCC TGTACCTGTT TAGGTCTTTC CTGGAGAAAT ATAGACCGGG AGATGTCATG
78781 GGACTCTTAA ATTCAATAGT GGTGGAGCAC TACACCAAAT GTCGTAGCGC GGAGTGTGTC
78841 ATTACAACCC ACGCCATGGT AGGATCTGGC GAAAACAACA AGGGACTTTT TTTCTTTCCA
78901 GTGTAATATA TAACCCATGT GTGTAAGTAG TTACGGTATA TTATTCACGG GCGTTTAAAT
78961 GCAATAACCC ACATAACAAA ATAAAAATGT GTAAAACCAA ACACGCGTCA AGCTTTTTCT
79021 GCAAGGGTTC TGTGCTCGGA ACAAATAGGC AGAGTTCGCC ACCTAGCGGC ACGTTCTCTA
79081 TAGCCTTGGT ATACAGCGGA TGTTATTTAC GGAAGTGCCA CATGAAACCG TGGCCGGGTG
79141 CGTAACAGTA AACCGGGTCC AATCGCGCAG TTTACACAAC ACATAAAATA GGCGACCGTT
79201 TATAGCGGTT TATTTGAACA TACAACACAC ATTTCAGGAG TTTCGGTCTT TATAAAAATC
79261 TTCATTCATT CAAAGTGCCG ATATATTTCA CGCTGGTGGA CAACAATACA TATCGGCGCC
79321 TGTGATTGAT CCTCCTCCTC GTATAGGGAC CAGGCAGACA CCCAAATCTG TGCAAACGGA
79381 TCGGGTGGCC CGGGATCAGA GGGTCGTGGT GACTGGGCTA ACTTTAACAT GTATTCATCG
79441 GTATCAAAAA CGCGATGTCC CACCCCAGAT GGCACGCGTC TTGGTTCAGA GGTGTTTGAG
79501 GTATTGCCAC GGAACCATAA ATTAAGAAAT CCATAGCCAT TCATCATAAT CCCATACTGA
79561 CAACTGCACA AAACAATGCC GTCCTCCAGG GCGTGTAACG CGTTCACAAG TCTGCTCAAC
79621 ATAAAATCGT CATCCATATA AGGCGCCGGG GGCAACCACA ACTGAAGCGG CCCGGGAGCC
79681 ATGCAGACGT GGTGTTCATT TTTATCCGTT GGTCTCGTGG ATAAACGAAC GCCCATCCCA
79741 GAATGGGTAA AAGTCTGTGC TTGAAGGACC CCATAATAAT AAACCTGAAT ATGCATAAAT
79801 CCCCACGGGT CGTCAACAGA TGGCGCCAAA CACGCAGAAA GACCGCTCGA TGCAGATAGC
79861 GGCAGGGCGT GTTCCCCGGG CAACGCAACC GCGCCCCCTT GACTATTGTC CTGCCTACCG
79921 TCTTCACCGG AGGCGGTTGC GGTTGCAGAC GCCTCGGTGG CCGGGGCGGT CATGGTGAAG
79981 GTACCGGCGG GACGGGTACG AAGTAGCCCG AGCAGCTGCT TTACCGTGTA TTCAGATACC
80041 CTGCCCGCCT CGTCGGCGCC CTTGCGCTTA AACATCCGAA ACCTCAAGCC CCGCAGTTCT
80101 GCCGCTACCC CAGAGGCCCG AGGGCACAAC CTGCATAACA TCTCTTCATC GGAGCGCAAC
80161 CTGAAGACCA CGTTTGCAAA TGAAAAACTC CGTCTCAAAA CGTCCTTTTC TTCTATGAAA
80221 TAGGCACTCT TTCTTATCGC ACCCAATAGG CGCGCCTTGG CCTGGCCCAG CCTAGACCTC
80281 CCGGACGGAA GGGGGATGTT TCTCTGGTTG CAGTAGTCCT CGTACATCTT CCCATCCCGA
80341 AGGTAATCAA ATCCCACAGA CCACGGGGTC GTAGTCGCCA GCACGACCCT GGTTTTTTCG
80401 TCGTCTGCCC AAAACAGTCC CGAATACGTT GCAAGATTTG CGTGAGTAAC GATCCATTCT
80461 CGAAGTCCCA TGGCGGACGG ACCACCGCTC TGAGGTGGGC GCCAGGTAGC CATGGCTGTA
80521 TATGCCAGGT TGATACTCAC CGCTTTCAGT TCGTGTTAAA ATAACCAAAC TCCGCCCATG
80581 CGGCTACCGC CGTGGGCGCA GATGTGTGCA GCCGTCACAA AAGAATCCCA CCTGCGGCAA
80641 CTAATTCCCT CAGCGCCCAG TCAGCTGGCT ACCGGTTCAT TTCACTCACG AGCACGCCAC

80701 CCCATCATCT TCACCGCTCG TTCCCGCGTC CGGATCGTCC ATTGAATCTG GTCCGGAGTC
80761 GCCCAGCTCT ACTGCCGGGT CATGCGATTG GCTCGTCCCC TCTCCTGGAT CTTCCATCTG
80821 GAATGAACAG TCGGTGCGAC CCTCTGATGT TACAGTCGTC CCGGGGTAAA AACAACCCTC
80881 TGTACCGCAT ACAGACGAGG TTTCGCACAC GTAACGCAGT CTCATGGCGA TGGGCGCATC
80941 GACGGGATTC GGCGACTGGA CCGACAGACA CAGGGCGATC TGTGCATAAT CTCTAGGTGG
81001 TTCGTCTCCA GGGTTCGGGG ACCTTGCAAG AGCTCTAAGA TAATCCACCA TATCAAAAAC
81061 CTGAACGGGT TGCCTGGATG GCAGGTCCTT ATACTCCTCC TCATTGTTAG TAAGCACCTT
81121 CATTTCCCGG TTCTCCAGGT TCTTCACGCA AATGCCAGTA TCCGTAAGGG TAATCAACAC
81181 ACCCTTCTCG CACTTTTTGT GCATAACCCT TAGATCTTCT CTGGCAAACT GACAGTCAAC
81241 GACCCCCGGT GACGGTAAAA ACAAAGTCCT CTCATCTTGG AAACAACACG CGTGGCCAGC
81301 GAGGACGGGA CATGGTCGAC TGCAGACCCG TATTCCAGCA TGAACACGTT CGAGGCCAAC
81361 GGTCTCTCCA AAATATAGAA ACGTTAATTG GATGGCCCCT GGCAGCCGAT GCTCGCGCAC
81421 ACGCTTGTTT TTGTCCATCC GACTCCGAGC CAATACCGTT GCGGTAAAAA CCCGTCTTCT
81481 TTCACGTCCA GACCCACCCC CCTGGTCGTA TTGGAACTCA CTGATTAAAC CAAGGCACTG
81541 AGTTTCGGTG GAAGACATCA GATTACACAT TGCACAACTT ACCACCGGTT CCGGCAACAG
81601 GCGAAACAGT CTATATCGAG CACCCGGTCG CCCATAGGCC TTATTAATAA ATTTAAGCTC
81661 TCGTATACTT TTGTGGTGTC TCATGTCGTA CAAGAATCTC ATTTTCCTGA ATCTGCCATA
81721 GTGGGACGGT TTTTCGCCCT GACATATGCC TCGCATGTTG CAGTAGTCAT CAAAAAACTT
81781 CTCGCTATTC CATACGGACC GATCTGCAGT TACTTTGTTC CACGGGACCC GGACAACGGT
81841 CTTGTCCTCA TTCTCCCAAA CCAATCCGCG GTACTCGCCC GATTCGACCG CGGCAACGAG
81901 CCAGGCCCTG ATATCGACTC CACGGCCCGC CATTCCGCGT ATTGGGAGCT TGACGGATCA
81961 GCTCTCAAGA CAAATGAAAC CCTTGCGTTC ACGGCACCAG TTCATAAATG TATCTAAATG
82021 ACGGAGTCAT GTGCCCTAAA CATATGGCTA AGACTACTCA AACTGTGCTT AGGGGAAGAG
82081 TGACAAACAC CTGGCACTCA ATAAAATTTT GAGGTCGGGC ATTGACCCGT GCCTTCCTCT
82141 GGTTAGTATT AACCCCATCG CCACCATATT CAAGGAACAA AACACTCCAG ATACAAAACA
82201 CTTTATTACA GATAGAAGGC GCTCACCGGC CTCACAAAAC CGGTCGCGCG CACAGGTAAC
82261 ATTAAACATT AGCGGCGGAC AACGCCCGCC ACAGACACTC GTGCCACAGC TGAATAATTA
82321 GCGGTACAGT GCTTTGCACT CCGCCCTCGG GTGTAGGTAC GCCCCCAAAA TATAGGTAAA
82381 CACAAGCTTG TGGAAGCGCC GCTCCGTCCG GAGATGTTCT GGCCATGGCT GTCATATAGT
82441 CAACCAAGTT AAATGCACGA CACGCGTTGG CGCACGACAG GACCGCAACC GCCCCCGTAG
82501 ATTCGACGTT CCCTTGAACG AACGCCAGGG AATGCCCCAA CCATCTAATG TAAATGCCGG
82561 AGGGGTCGGC CCAAAACGCC AACCCGCGCA AAAGATCTTT ATCCAGGAAG CGCAACGCCT
82621 CGGCCTGAAA CTCTGCTAAG TGGCCTGGAA TCTCCGGAAC CAGTGCCTGC TCGGCGATCG
82681 GTGCGGCGCA TATGTGCCCT TGAGGGCGAC GTTCGGACAA TGGGAGCACA CGGATGCCGT
82741 TGGACGATTC AGTCACCACT TCACCAACTC TGTCCCCATA ATAGTAAAAC ATTATACGTA
82801 GCATCCAACA CTGTTCACCC CCGGCTCCCA CCGCACCAGT GTACCGAACA CCGGCGCCAA
82861 CCGGACCAAA CTTATTCACG GCCTCCTCGC GCAACCCCCG TAACATAGCA GTAGTGCCGG
82921 AGGCCTGATC GCAACGAGCA CACGACCTCA CGGTAATTGG TAAAAGGCGA ATGATTCGGC
82981 ATCTGTCCGG ATATGGAGTC GAAAGGTTGT CCTTAGTTGA CACGTCCGAC ACGGTTTGAC
83041 TATGCCGTAT GGCGCTGGAT AGCCAGTTTT TGCATTCCTT AGCGGTTAAT TCCCTGCCAG
83101 CCGCGTGCAA GATGCCCCTG GATCGACAGT AGTCAATAAA AATGTTTTTT TCCGCCTCCT
83161 CGACTCCTCT GCTCCCTCTG TCGTGATTCC ACGGTATGCG GATTAAAGTT CTTTCCTCGT
83221 CAACCCACCG CATTCCGGGA TGACGACCTG TCTCGCAGCA TTCCACCAGC CATCCTCGAA
83281 GGCTGGATGG TTTGGTCACT CGGACCGGCC GCTCCATATC CTCGCCGCGT GCACGGCAAT
83341 GCCAAACCTT GTCCAAGTTA CCTCCTGACA AGCACCTGCG ACCCTCACAG ACCTACGCGT
83401 GGCAAACGGG AACTTGTGGT CTAGTCTGCT GGGACAGTAC CTTAGCGTTA TTTTTATCTG
83461 TCGTGCACTC CTGCGGCGGA AGTTTGTGGT TTATGAGCAC CACCGGGAAG GATTCATGTG
83521 TCCCTGCACC AGCTCGTCCT GACAGACCAC CACAGAAATA GGAACTTGTG GGCTCTCTTG
83581 GGCACCAGGC TGCTGTGAGA GATACAGTGT TGCCCCGTTG CACGGCGCAT GGCTCTCTGG
83641 TCCCGGAGAG CGGGCAAGCT TTCTCATGAA ACCCATGAAG TTAAAAATTT TCTGTTTAGA
83701 GAGGAACATG ACCTCTCGTT GGATAATATC ACCATTACAG TCGTTACCCG CCGCGTACAG
83761 CGTTTCCCTG TTCCTATTAT CGACGTACAC TCCGGTTTCG TTGCCGTGTA CGTAGATGCC
83821 CTCGTTATAC GCGACCAGCA TCGTTTTTAT TTGCGTTATC TGTTCTGCGG TCAGTTTTTC
83881 ACGATCTGGA GATGGAAACC AGACAGTGAA CCCGGTGCCA TAGAAACACA GATGACCGGG
83941 ATGGGGAACC GGGCGCGGAT GCAAACGAAC ACCTCTCCGA TCTTGGGACG TGTGATCCAG
84001 TGCCAGAACG CCAAAATAAA ACACCTGAAT TCTCAACCGC CAAAACGGCA CAATTTGATG
84061 AGGCAGCGGC GCTTCTGGGA CCCCGAGACG ATGTGCTCGG GCGTACGCTG ATCTCGTCGG

84121 AGGGACCAAC GTGGGCAAGG GGGCGGGCAG CGGCATTCCC GGCTCTAACC CGATGCCACA
84181 CACCTGCTCC AAATAATTTA TAATCCCTTC AGAGTTCTCA TCAAGGATAC ACGCATAGCA
84241 TCCAGAATCC ATCAAAGGTT TTACCCGGAA GGCGGTCCAG CCGCCACCGT GCGTCCGCAC
84301 GACCCCATGC TCCACTTCTT GAAATCCAGC ATTCTCCCTC AAACCCCTCA AAAATCGTCT
84361 CTTACACTCT AGCAGGTTTC CATCTAGCCC AACCTGCAGA CCTCTCTCCA CGCAGTAGGC
84421 GACAACGGCT GGAGTTCCAG CTGCCCCCTT CTTCCAAGAC AGCTTAAAGC TTTTTTTCCC
84481 TCATCGCACC AAGTTAGGTC TGAATAGGTT TTCTCGTTGA GATGGTAAAG GGTCCACGAT
84541 CGTAGATGGC TGTAGTTAGA ATCTCGATTC GCCATGACAG CGTACACCTA TACAATAAGG
84601 CGCGATACGA GTGTACTTTT CCCCTGATGA TATTTGCAGG GCTCTATCTC CTCAAAACGT
84661 GCACCTGGCT AAGCAAGCAG CTCTATTTCT GCCACTTCCT CATTTATAAA ACCATATTGT
84721 TTCTGGTCTC ACGTTTTGCA GCTCACCTGC AACGACACGG TGAATGATCC CGAGTTTCGC
84781 TCCTCCAGTG CCTTGACAAA CACTGGCCCA GAGTTAAAAA GTGAAGCAAA AGGCATAGCT
84841 TTCGATTTCA GTCACCGGCA GTGAGTGGTG GACATTAGAA AGTGTTTGGC CCACACATTC
84901 AGCTGTGCAC TCAAGCCCGA CCAATATGGC ACGTGACCAT TCCCCTATCA CAACACTAGT
84961 ACAGAAAAAA CGAAACCACA CGCGAGGTGA CATTTTCCAG GTTAGAGAAA TTTATTAAGC
85021 CGGCGAATCC ATCTAACAAA ACAGCAAATG TTTATTCAAA GTGCCTATAG ATTTCTGTTT
85081 GATAAACAAT AATAGATAGT GGAGCCCTCG AAGAGTTTCG CTCCTCGCCC AGGGAACAAC
85141 CCGAAACCCA GAGTTTAACA AAGGCCGCGG GGGGACCCGT GTCGCGCGGG GAGGGGGACC
85201 TGGCCATATC AAGCAAATAA CGATCAGTGT CAAATGCCCA CACGTGCATC CCGGTGTAGG
85261 CACGATGTGG CAAACTGTAC GGTTCCACGG TATTTCCCTG AAACCACACG TTAGGACCCC
85321 CGGTAATCGT GAACATAATT CCCCTGCTGG TGCTGCAGAA AATTATTCCA TCCACGAGCC
85381 ATTTCAAAGC GTCCTCGAGT CTCTGCACCA CAGACGACTC AGTCGTCTGT GGTATCGGAG
85441 GAAACCACGC CTGAAGGGGT CCATCCAAAA CACACACGTG GTGTCCCGCC TTATTTGTTG
85501 GTTTTGAAGA TAGCCTGATT CCCTGCCCGG TGGCTGTCAA CAGCTCAGCC TGCAATTGCC
85561 CATAGTAATA AATCTTTATA TGCATGTGAC CCCATGGATC GCCGGCAGAT GGCGCAGCAC
85621 ACGCCATAAG TCTGGCGCGC AGAAGGCCCG AAGAATACGC TGCATCGTCG TCCTGCGCGT
85681 CGTGTCCATC TTGATGACCT TCCCGATACC GTCTCTCTTG ATCACTTACG GACGGTCCCG
85741 AAAGGGCGTG AAGCATTCCG CGCCCGAGCA TGGCAAACCG CATATTGCGC AACTCAAGGC
85801 GGAGCGCGCA CACGCGCGGA CACACCGGAC AAGACATCTC TTCGGATGAA CGCAACCTAA
85861 AGACCAGCGC CGTGAACGGA AATTGATCGG TAGGGAATTC CTTATCTTCC TCAAAGTACT
85921 TGCTCTTTCT TACGGCTCCC AGCAGGCGCC CCCTGGCTTG ACACAGCCTG CTCCTTCCAG
85981 ACGGCAGTGG GATGTGTCTT AACTCACAGT AGGCATCATA GTGTTGACCG TCGCGTTGAT
86041 AATCATAATT GGGAAACGAC GGTGGTGTAG CTGCCAACAC CAGCCTCGTC TTTTGCTCGT
86101 CTGCCCAAAA CAGGCCCGGG TACAGACCAC TATCGGCTTT GAACGTCAGC CATTCCTTGA
86161 GGGCCATGCC ATAACGAGAT GGGCCCCTTC TCGATTCCCC AGCGGCCATG ATTCAGTCGG
86221 CAACGTTTAA AACCCATGTA AAGTTTCAGT TCAATTTAAA ATGACACAAC TCCGCCCATG
86281 CTGTGACAAA CATAAAGAAA GTGGCACACG TGTGCGCTTT GGCCGTCCTG TCAGATGAAC
86341 CCCGCCGGTA ACATTTACTT CCTCCCATAC GGTGCGAGGG TAGATACAGA TGTGCTTCCC
86401 GTTGTTCCCG AAGCACCGAC ACCCGCTCCA GACGGTCCCT CGTCCCCGCT GTCATCTCCC
86461 GCGGCACCGT TGCTGACATC ATCTGACACA TCATCATTGC ACACGTAGCG AAGTTTCAGG
86521 GCAATGGGAG CATCCCATGT ATTTTCGGAC TGAGTAGACG GGCACATGAA AATCCACGCA
86581 CAGTCACGCG GCGGGACATC TCCTGGAGTT GGCGATCTCG CCAGCTCCCT CAAATAATGC
86641 AAGAGATCGA ACGCCTGGAG CGGCTGCGAA GGCCGCAGCT CATAGTAATT TTCAGAATAG
86701 TTGGTCAATG CCGTCATGTT GCGTTCTTCC AGGTTTTTCA CACAGATTCC AGACTCCGTC
86761 ATAACACAGA TCAAACCTCT CTCGCATGTT TTCTGCAGGG CCTGCAAATC TGCCCTGGCA
86821 AACTCGCAGG GAATCAACTG GGACGACGGC AAGAAACACG TCAATGGATT ATCGGCACAA
86881 CACAGGTGTC CGGGACGTTT GGGATCAGGA AGGCCACATA TACGCACACC CGAGCCGGCT
86941 CGCAAAATTT GAACTCCGAC GTTCTCTCCA AAGTAAAAAA AGGATACCTC AACAGCCCCT
87001 AGCACAGGTT GCGCAGAAGA GTCCTGGAGT CGTTTGGATT TCCGTTTATA AACGGAGACT
87061 CTCAGTGGCG CCTTAGCCCT CGGCCGCCGT GTCCTCTCGA AAACATCCTC TCGAACGCAA
87121 CTGCCCATGG AACAGCTATG CAACGATGTG GTCGCGATGA GGTTACACAG ATCGCAGCCC
87181 ACCGTAACCT CAGGAAGCAA TTGAAAGAGT CTGTATCTTT CCCCAGAGAC GCCCCCGGCC
87241 TTGCTCGGCG GTGTCAGCTC CCTCAAATAC CTGTGAGATC TCACTGCATA TAACATCCTA
87301 ATTTTCTTAA ATTTCTGGGC ATGTGAACCG GTTTGACACA CGCCACGCGC TGAGCAAAAA
87361 TCGTCAAATA TCTTGTTATA ATCTTCATCG ACGCGACTGT CGGTGCACCG GTTCCACGGG
87421 ACACGTATAA TGGTCTTGTC CTCGTCATCC CATTCCACTC CGGGGTATCG CTTAGACTCC
87481 ACGGCCTCGA TGAACCAAGC TTTTAAATCC ATATCGCGTT CCGCCATCTC GCGTATTCAA

87541 TACCTCTGAC CGAGCGGCTC AGTTAAAATA CCGAGCCCCC AGAACCCCAG GCCTCCGTAA
87601 CAAAAAGGAT CTAGATCTGA AAGAAGCCTA ACTATATACG CCCACGGGAA AACTCTGTAA
87661 CAAATGAAAA AAATGTTCTC AGAAAACGAA GAGAGGAAAA CTGACGACAC GCAATCTAGT
87721 CAAACCAGAA CCTAAAGGCC GTCACTGACC CATATCTTTC CCCTTCCCGA GCGAGTGCGG
87781 TTTCTCACAT GCCGTGAGCA AACGAAAACG CTACACATAA TAAGACACGT GTGAGAGGAA
87841 AACTTTATTG CAGGGACAGG GCAAAAGCAA GCTGTGCACG GTAACAGTAT GTGTCACTGG
87901 GCCGCATCCC CGCACGCCCG CCACAGACAC TCGTGCCAAA TTTGAATGGA CAGGGGGATT
87961 GTGGGACACG GCTCGCCTTC CGGAGTGGGA ACGCCTCCTA GGTGGAGAGA CGCATAGGCG
88021 GCTGGAAATC GCGTTCCGTG TGGCGGGGTG GCTGACACCG CACTCAGATA TTGGGCGAGC
88081 GAAAACACGC GCAGAACTCG CGCGTTGTTA ATTATATCCA CCGGCTCCAG CGAACCTCCG
88141 TTACCCAGCA GGAAGGCTGG AACGTGACCC ATATACCGAA TGTAAATCCC AGACCCGTCT
88201 GCGTAAATCA TCAACCCGCG TATGATCTCC TTCCCCAGCC ACTTTAGCAT GCTCGTCTGG
88261 TGAGGAAACA TATCCTGTGG GGTATGCGGC AACAGGGCCT GTTCCGGCCG AATGCCCGCA
88321 CAAACGTGTC CTTGCAGTGG TCGCTTTGGC GCCGGCAACA GGCGAACGCC GTTGGGCGAC
88381 CCGACCTCCA TGTTCCCAAC CAACCGCCCA TAGTAATAAA CAACCAGCCG TAGCATCCAA
88441 CACTGCTCCC CGTCGCCTCC CACGGCGCCT TCGTACTCCA TCCCGGCACC GAGATACCGG
88501 AATCTGGCGG TTACCTCGTT GCGTACGTCC AGGAGCATGC CGCCGGTGGC ATCCGCTTCG
88561 TTACAGAGCG GGCAAGATCG TACAAAGATC GGCAACAGGC GAATAACGCG GCACCTGTTA
88621 GGCGCCGGCG CCGACAGGTT CTCCTCAGTG GATACATCTT CGACGGTCTG ACTGTGGCGT
88681 ATAGCGCTGG CCAACCAGTT CTTACATTCC CTGACGCTCA ATCGTCTGTT ACCGGTCTGC
88741 CGTATGCCTT TAAATTGACA ATAGTCCAGA TACATGTTTC TCTCATCGTC AGAGACGCCA
88801 CCCGCGCCCT TTAAATGATT CCAAGGTAAA CGAACCAGAG TCTTTTCCTC ATCCAGCCAA
88861 TGCATTCCAG GATGCTTATC ATTATCGCAA CAATCGAGCA ACCATGCCCT CAGGCCAGAC
88921 GGCCGGTTTA CGCGGATCGA CCCTGCCCGT CCTTCCGCCA TCTCTGAGGG TCCGCTCGCA
88981 TAAACTGGCT TAAAAATCTA TCCCCCCCCG TGCACCTGTA CTTTACAGTC ACACCTTCGG
89041 GCACACGCCG TAGGGCTGGC AGGGACGGCA CCCAGCAGCT ATTTTCATCT CTAGCGCCGC
89101 CTACTAGCGG AAATCTGCGG TCTACAAAAT CCACCGAGCC GCCCGAATGT CACCATGAGT
89161 CAATTCATCC TGACAGATCA CCAAAGTGAC AGGGACCTGG GGTCTTTTAC TGGAGCCAGG
89221 GCGACCCCCA AGGTAAAGAG TGACTGCGTT GCTGGGAACA GTGTCGCCGG GAACGGGTGA
89281 CCGGGCGAGT CTGCGCAGTA GACCCATAAC ATAGAAGATT TGAGATTTGA CAGAGGATCG
89341 CACAGCCCGT TGTGCGTGGT TCCCCTCTGC GTCGTTCCCT GCGCAGCGAA CCGTGTGCCT
89401 GGTTCTGTTA TCCACATACA TCCCCAGCTC ATTGCCATGC AAAAAGACGC CCTCCTCACA
89461 CCCGTTTATA ATCTCACACA CGGTTTCTCT CTGCTCCGGC GTTAGCGGCC CATCCAACTG
89521 AGGTTTTGGA AGCCACATCT TATACGCTGT CCCATAGAAA CACTCGTGAC CCGGTTTGGG
89581 GTCTTGGCGC TTATGCAATC GTACCCCTCT ACGGTCCTGG GACGTGTGTT CCGCTACCAG
89641 CGACCCAAAA TAAAAAACTC GTATACGTAG CCGCCAAAAT GGAGTGATCT GGTCTGGATG
89701 CGGTGGTGCA ACCGTCGCCA ACCTGGCGGC CCTGGCATAT ACGGACCGCG CCGCACCGTC
89761 AGCCTCCGCC GGAGCGGGAG GAGGCAGGGG CGTCCCTGGC TCAAGCCCAT GGACGCAGGC
89821 CGCCTCTAAG TAATTTAGGA CGGTCTCCGA ATGACTATCT AAGATGCAGG CATAACACAG
89881 CGGATCCACG GCGGGTCTCA AGTAAAAAAA TCTTTCTCCG CCAAAACGCG TTGTTGTTGC
89941 CCGGGCATCG TCCTGTTCAA ACCCAGCGTT TTCCCGAAGC ACCCTTAGCA GTCTTTTCCG
90001 GCACTCCGAC ACGTTGTGCT GGCGTCCACA TTCCAGATCC CTATCCAGGC AGTACGCAAC
90061 CACAGGTTGC ATGCCACTCA TACCTCGGTG CCATGATATC CTAAAGCTCC TCTTTTCCTC
90121 GTCGCACCAA CACAAGTGTT CGGGAAACGT GTTCGCTTCC AAATTGGAAA TAATCCATCT
90181 TCGCAGGTGA TTGTGAGTGA TTTCTATCTC GGTCATTGTG AACTTGGTCC ACGAAAACTA
90241 ACTCAAAAGC CCTCACCACT CCCCTAAACT GTAGCTGAAT TTTATGTCCA CCTTATAGGG
90301 TGCACGTGTA CGTATATGCC CCGTGAAAAC CTCGCACTTC CTCATAAATT ATCCCCAAAG
90361 CGTTTCGGGT CGCTCACCTA AACTAACTCG AGCATTACAC GGTTCATAGT GTCCCGGGTT
90421 TCGCCACTCC CCCACCGCGA CAGAGAAACG AAACGTTCAA TTTAAAGGGC ACCAAGCTCA
90481 ACTTTATTTA AGGAGAAAAA AACCACCGCC ATCAGAAGGT TTCCGCAGGC ACACACCCCC
90541 AAGACTATTC TTGGCGAGTT CGCGGCACGC TGTAGTTTTT TTCTGATTAT TTGTAATTCG
90601 CATCCAATCG CCCCAAACAT AGTCAATAAA AAAACCGTAA ACACACACAC TCTGGTGAAC
90661 AAATTATGCG TTCCTGGCCC CAAGACGTGC GTGAGCGTCA GCAGCACAGC GGCTGCTAAA
90721 ACATAAACCA GGGCTAAAAC GTTTGACCGA AACACGAGTC CGAGGCAACA AAAAAGCCCA
90781 ACGTGGAAGA CAAAAAAATA GTGTACAAGT CCGAGGAGCA GAGGCGTCGT CAGGCCAACA
90841 GACAGCGTAC CCGCGGCCGC CCCTCCAATC AAAACAGTCA CGAGATAAAA GTCCCATCCA
90901 CACACCTCAA GCGCTTGCTT TATAGAGTAT GTGATGTACC GCCTAGAGGT GAGCGAAAAA

90961 ATTGCATGGC GTCGCTGCAG GCCCGCCTCA TAAAGCGGAT GTGACCGCAG CGCTCTGAGT
91021 GAACAAACGC CACCTACTAT AAAGGCCACC AGGGGCAGCA AAAACACCGG GGCTAAAAAA
91081 AGGTTCTTAA GAACTAAAAA ATAATACACC GAGAACGCTG CTAAAAATCC AAGTTCGTAG
91141 TATCTGTGGG CAATCGTGGG GCACAAATAT ACCACGTCAT TCGCCGCAAC AACACAGAGC
91201 AGCAGGCACA GGCACGACCC CACAAAAATG GAAACGTGAA CCCATGGAAG GTCAACGCCG
91261 AGGGTGGATG CCATAGAACC GGCGGTAATA AGAAACGCCG AGATGGTGCA CGGTACATCC
91321 ACCTGCCGCG AGAGCCAGGC CGGTAGCCCA AATATACAGA ATAGCAAAAG CAGCCCATGT
91381 AAGCCAACCG CGCCCCAGGA GAATATAATG TTCTCAACGC TAATCAGGGG CTCTACAAAA
91441 TACAGCGAGT GATAAATAAA ACACAGAATA AACGGGGAAG CCGAGACCGT CCCGCTCATA
91501 AAAGATAGCC AGGCGAGGGA CGCCTCCGAC GTGTACGTGC CCATGGCGCG GCGTTTAAAA
91561 CAACGGGTTG AACGTCGCGC GAGGCGTCTT GGTTTTTCCT TTCCCGCGAC TGGGTGAGCT
91621 GCTGTGTTTT CGCTTGCCGT GTTTTTTTTC CGGTCGAGAG CCGCTCGTCC GCAGTGATTT
91681 GGCGCCCCCG TCTGTCGGCT CGGCACCGTC CTCGGCAATT GGGACGCACT GCTCTTTTAA
91741 GTCCAGGTCA GGCGTTCTCG GAGGGGACGG GAGCTCAGGA ACCCCCTCGG GATCGGCACC
91801 TTCCGAGAGA TTGCCCGGGA CGGCATTATT CCGGTCCGTC GGATCTACCT TTAATAGCAT
91861 GACGCCCAAC CTGACGTCAG CAAGGCTATC TGTCAGCGGC AGGCCCTCAA CTTCCAGAAT
91921 CTCATTGGCG TGCCAGCGGA AACCAGGCAA AAAAAATCCC GGAATCTTAC ATGCGTTCAA
91981 TGCGGCCACC AGAGACTCTA AACTCACATG AGTCGTCAGG TCGCTAGAAA CCACACCCAC
92041 GTCCAGCTTT CTTCCTCTAG CGCAAGTCCA CTTGGCGTCC GGGGTGTCGG GTTTAAAATC
92101 AATAATTTTA CTAACCCCGT CGACCGTAAA CGTGGCAACT GCCAGCGTTT CGCTTAACGA
92161 GACGGTGACC ACCTGCTTGG CCCTGGTTTT CTGCTGCCTT AGCCACTTTA ACAACAAAGT
92221 CAGCGTCTTA ACCGTAAGGA ACAGGCGATT CCCCGTCCCG GCCTGCTCTA GCCTGTCGTG
92281 GAACGAGGGT AGGTTGGGTT CGTCCACGAC GCTCTTGTGC ACGGTGCTCG TGTGATTGTC
92341 CCCGTAGGCG ATGGTCGTCC TAACAAACTG CGGCTTTCCT TCCGCGTCCG ACTCGTCCCC
92401 GTGGTGGTAG AACGCCACGT TCATTTCCGC AATGTTAACT CCAAACATCT CACGGGTGTG
92461 TGTGAACGTG TTCCCAAGGG AGGTGTTTCG AAAGCTAAAG CTAACCTCGT CCGTCATGTC
92521 AGAATATACC ATCAGCGGCG CCAGTGCATT TTTAAGCCTG AGACCGAGAA CGCCCGCGTC
92581 CCCCACGCTG GAGAGCACGC TAAGGACGGG GGCCCTGCCC TGGCCGCTGA TCTGTATAAC
92641 GCCTTTCTTA ACGATGCCCT TAATGTGGTC GTAAACCCTG CTAATGCTGC CAAGCGCATC
92701 GACGTCCACC CTGGCGCCGT AATGAAACGA GACAGGCATG GCTAAGGCGA AATTAACACC
92761 CTAAAAGACG AGAAGAGAAT GGCAAACACG ACAGACACAG CGTTCAATGT TCAACCTTTT
92821 AATAGAGCCT CTGGCCCGGA AAATGAGTAA CATGGGACAC ATGGGGACTA CAGATCATTA
92881 TGAACGGATG TGACATATTC AGAATTATCA CGCTCAAAAA AGTTCACACT CTTGGCACTC
92941 GTATACGACA AGGGACACGC GTTTGGGGGG CGGGCACCGA ACAGGGGCGG TATGTTTATT
93001 GACTTTAGAA TCCTGTCGGC CGTGGCCTGC AAAAACTGCT CGATGTCTCC GACGTTCACC
93061 ATGGTCACGT CTCGGCTCTT GGAACGAATA AACAGAGTCT CTATCTCCAC GGCTTCGCGA
93121 AAGAGGCCGT GGATGTATGC GACGGAGGGG GATTCGTCTC GGGAGATCAT GGTGTTGTAC
93181 AGCAAGCTGG CGGCCCTGGT GTGTAGCAGC TCGTCCCTGG AGATGTAGTC GTTAGCCAGG
93241 CACACCCCGC GCATGATTCC GCGAAGGCGA AAAAGTCCTA TAGAGTAAAA TGAACTTATA
93301 AAGTATATCC CTTCGATAAC CAGGAACAGC AGCACCTTCT CGGCCCGCGT GGTGGCCTTG
93361 CGCACTCGGG AATGCAGCCA GTCAATCTTC TTCGCCAGGG TAGCGTCCTC GACTATTTTT
93421 TTGGCATAAA TCATTAAGTC CCCCCGATTA CCGCCAAAAA ATAAGTTCAG TATGTTGGCA
93481 TACACTTTGC CGTGGATGTT TTCCATGGCT ACCTGTTCGG CGTAATAGTG GGTCACGTCG
93541 TGATTGCTGA ACTCTTTAAC TAGGTCTTCG ATGTTAAAGT TTACCAGACT CTCTGCCATA
93601 CCCAAAAATG TGAAGAGAAA CTGGTAAAAC TCCCTATCGG TGGGACTCAA TAGTCGCACC
93661 TCATCGGAGT CCTCGTGCAG GGAAATTTGA CTGGGGAACC ATCTGTTCTG CCATGTCTCC
93721 TGGGTGAGTG CTAAAAAGCC AGGGTGGTCG CTCGCGTACA GAAACCGATC TGGGGACTCC
93781 ATGGCGGCGG TTACTATTGA CAGGCCAAAC ATACCTTGGT CCGCTCACCG TCGGCCTTTT
93841 GTACGCAGTT AGACTCCGCC CCCTCTGAAG AGATGTTGTC ACCCACGACC TGATCTAAGC
93901 ACTGCAGGGC ATTAAGTTTT GTCTGTTTCT GAATTCGACA ATAATACAGC ATGGTTTTTA
93961 ACCCTAGCCT GTATCCGGTC AGGAGCAAGT CTCGAATATA ACTCGCCCGC GTGGCCTGCT
94021 CCTCCTTTAA AAAGAAATTA AGCGACTGAC TCTGGTCTAC AAATGGAGCC CGCGCACCGG
94081 CCCGCTCCAA CTGCTTAATT GGGCAGTAAT CAAAGGCCGT CAAAAAAATC TTGTATCTGT
94141 CCTTCAGGGG TTCTGGAAAT GAAGCCACGT CACCTCCGTA ACGGCGCACC GTTCGCAGGT
94201 CACCAGGCTT CACTCTCTTA AAAAAAGTCA CGTTTGGCTT CAAGATCTCT TCTTTGCTGG
94261 TGACCTTCGA TGCGATATTG GCAAAGAACG GATAAAATGC TTCGGTGTAC CCCGTAAGCT
94321 GAGAGGTTCC GGCCGTTGGC ATCAGCGCCA AAAACTGGCT GTTGAAAATT CCATGCTGGG

94381 CAATGCTGCG CCCCAGCTGC TCCCATCTCT CCAAGGGAGG GTGGGACGGC TTAACGCCGT
94441 CCCACGTTTG CCAGTGAAAC ACACCCTGAG CCAATCGACT CCGCTCCCAT CCACGAAACG
94501 GAGTCCCCTC GCCGAGTAAA ACAATCTCAT GACTGGTTTG CACCGCCGTA AAATACATTG
94561 CCTGAAAAAT CTCCACGTCC AGCTTGGCGC TCTCGGCGTC AAGGTAGCCA AAGCCCAGTT
94621 CGGCAAACAC GTCGGCCAGT CCTTGGACGC CAATGCCCAT TGATCGTTCC TCCTGGCCGC
94681 GCCTAACGCT CTCGGTAGGC GCCGTCCCGC CCAAAATGCA CGCATTGACA ATGATAACTG
94741 CGGCCTCCAC GGCATCATCG AGCAGTTCAA AACCAAATGT TACATCCCCC TTCCCAGGCT
94801 CCACACCAGA CTCTCCCCTG TGGGGTCTGA GGCACTTTGG TAGGCTAATG TTTGCCAAGT
94861 TACACACGGA AGCCTGACCC TCGGGTTGCT GCACGATTTC CGCACAGAGA TTAGAACAGT
94921 TTATGGCGCT GCCTTGCGTC TCACACCAGT GGTGTTTGTT GAGCGCCTCC TTTAGCAGGA
94981 CGTAGGGACT GCCGGTCTTA ATGACAGTGT TAATAAGGGC ATACATCATC GATTTTAACG
95041 GCAACGAGCT AGAGTGTTTG CCAGCGGCCA CTAGTCTGTT GTATTCAATC TCAAATTCGG
95101 CACCGTAGAG TTTTAGGAGA TTTGGCGCCA CCTCTGGCGC AAACAGGTGC CACTGGCCAT
95161 CTGGGTTTGT TTCGTACAGT CGGAAAAAAA GCTCCGGCAC ACACACGCCC TGAAACAGGT
95221 TGTGACACCG CTCCTGATTC TCCGGCATCT TCGCGTTCAA AAAATCACAA ATCTGATGAT
95281 GCCATAGTTC CATGTAGGCG CTGGCACCAA CAGGACGGAT ATTGTTATCG TTGAAATACC
95341 CAACGTGGGC GTTTATTAAT TTTAAACAGC TGGTGATGTT CTTGTGTTCC GCAAAAGACG
95401 AGACATCTAT CCCCACGCCT GACTTGCTAG CGAGAAGGGG GGACATTTCC TCATGAAGTG
95461 CTTTGAGGGT TTTGTCTTCG GTCGCCATGG ACGGCTTTAA AATAAAACAG CTAGAAAGCT
95521 GACCTCCGCG AAGCCCGGCT GACCTTAATA CAGGCGTTGC GCAGCACACA ATCTGTGACG
95581 AGATGTAATG GAACGCGTAA CCAACCAGGT ACATCTCATC CAGCTCCGTT TCGCTCTCCA
95641 CCAGGTGTCT GAGGGTCTCC CGCAAACACG GAAATTTTAT ACACTGACAG GCCACAAAAA
95701 CAGCCACCCT CATAAACATC TGGGCCACGC TTTCAAAAAT GGGTGAAGAA CCCTGGGTTC
95761 TCAGCACGTA CGTATCGTAA AACCTAACGG CAGACAGGTA GCCGCAGTTA ACGAAATTTG
95821 TGTATGCCTT GCTCTGCTTA AAGTCCTGTA AAAGACCGTC AAGGGCCGCT TCGTGTTTTG
95881 ACATAAACGC GCGAACCTCG TCGCTTAGCC TTTCCCCGAA TACCGCGAGA TAGTCTCGCA
95941 CCGTAACCAG GCACCTGTCC TCCATTATTC TGTGCCAAAG GAGACCGGTT AGTGAGTTGG
96001 CCTCTATGTC CCATCCAAGT GTTGCCTTTA AAGAATTCAC GAGTGATTCG GCGCATCGTT
96061 CAGGGTCAAG GCTTGAGCTA CACCCCCCAG TTCCCGCGTC GGTAACTAGG GTTAAAGGTT
96121 TGGCAGATTT TGCCGCAGAG AAAGAAGTTT CGGTGTTCAT TGTTAAAGAA ATAGCTTCCA
96181 AGACCCCGGA TTGGAAACAA ACTCCGTAAA TTTTAACACC GGTAAAGCAG CGCCTTTAAA
96241 GTGAAGGCTT TGAAAAGATG GTTGTAAACC GGAAGGCACG CTTCCAAGTC TGCAAACTAC
96301 GCCGAACGCA AGCCTATTTA TATACAGGTC ATTCTGCAGC TGAATGTATT TGGTGCGAAT
96361 CACGCCGCTG TAAAAATCCC TCAATTGGGC AGCTATTTCA CAATATCCTT TACCAGACTT
96421 AAAAAACCCA AGTCCTAGCG CCGGTCTATT ATTTACCCTA ATGTAGGTTA AGCAAACATA
96481 AATAGACGAG TGGGCGGGCA AAGGCTCGTC GCGCAGGCTG GACGGGCATG AGTCATCTAA
96541 ATCCACAAAC ATGTCACTAG GAAGCGTAAG GCCAATATGT GTAACAACGG GCTCTCTGGC
96601 GACTACGTTG CCCTTTAACG CAGACGTCAC CTTGGTGACA AACGTACTGT GGACCGTTTG
96661 AACCAACGGC CCGACCGGCG CAAGAAACTG ATGAAGCGAG CCGGTTTCCA ACAATTCTTC
96721 AAAATTGGGT ATGGCGTCAA GTAGACCGCT CTCGTGGCCG TACCAAACAC ACGCTATTCT
96781 GTTGGTCTGG GGGGCAGAGT CCGCGTCCAT CCTAGACAGT CGCGCCAGCG ACGTAGGCGT
96841 GAATAACATG TCAATGGAGG ACCCAGTGTC AGTCTGTTTA AAGGAAAACA GGTAGGTGCC
96901 CCGAGGTTCC TGTGAACTCA TGGTCTGAGA ATAAATCAAA AAATCTCCAT ACGTTTGACA
96961 TGTAGGCGAG ACAGATAAAA ATCCATCTTT GATGGCCTCC ACCCCAGTGG TGGTCGACAC
97021 CACATATTTA GAGAGCAGAT CACGAACACC CTTAGAAAAG TCGCGACCGC GAGATAACGA
97081 AACCCGGTGA GGAGGCGGCG GCAGTAGACG CATCAAACTA TCATTCAGCT TGTTCACGTT
97141 TGCGTCCCTG GTCTTCATGG CGTCGTCAAT TCCTGCCGCA CGGGCTGATA ACGGTGACGA
97201 GAACACCGGC GGGCTTTACA AATTAACTGA CAACCTCCTA ACGTGCACCG GATCGCTACA
97261 ACAGCTTAAA CTCCTGATGG AGTTCCAACT AAAACCACTA CCAACCGCAC ACCTTTTAAG
97321 CATGCCCACC GTGACCCGGT TTTTAAATAC TGCATTCAAA ATAGACAACC CCCTGGTTTC
97381 ATTTATTCAA AAACACCCTG TGTTTTTTTT AATGAGAGTC GCCAGGCTTC CGGAGCCAGT
97441 CATTACAGAC CACCAAAGCG CAGAAACGTC AACAGGCATA CTATCCGAGG TTGTGAATGT
97501 TCTTAATACA GCTATTCGAA AACCACACGA GTCCCCAGCG GCTAAAGACA ACGATTATCT
97561 CGACAACCGT GCCATATTGG CCATGATTAC AGAATACATC CATCACGTAA CTTCACGTAC
97621 GCCCTCGGGG ATCCCACCGA CACCACCAAT GGGTATCAGC CATCTACCGT GCGTAGAGCA
97681 AATTTTACAC GAAACCCACC GGCAATACTG GAACCTAACC CTCCCGGAGT CGCTATTTAT
97741 CGACATCGGC GAAGTCGCGT CTCCGCTTCA GACGTGGCTA ATCCTATCGT ATTGTAAAAA

97801 ACTGCAACTG GCACCACCGC CCCTATTCCC ACCTGTCGAC GAACTGGCCC GTCGTCTGGT
97861 TACCGGCCAC CACGAATTGT TTGTCCCCCT ATCGACGTCA CTGGAAACGT ACATCACCAT
97921 GCCAGTATCA AAACAGCGAG CGTTTGAGAT ATACAGCGTC TTTGCCAAAT CAAAAAACAT
97981 AGTTGACGGC ACACCAATTT TAGCGTTCAC CGACACAGAA CTTACAACGT TCACCCCCGA
98041 ACTGTTATTC CTGTACGACT TCGTAATCGA GTCCCTGTGC AAAAACCAGG CATACGGGTG
98101 TTCGCGCAAC GCAATTGAGC ATTTTATTAA GAAAGGTATC GATTTTATGG CGGAGTTGGG
98161 GGCGTTCATT GAAAAAACGT GTGGCTATCG ATCAACCGTA AGCCTGTCCA ACGTGAGAGC
98221 CGTTAAGGCC AGGCTCGCGA GCTGTGGCCT ATCCAAAGAG GCGTGTGAAG ATTTCCGCGC
98281 AATGATACTG ATGACACCGC ACGAGACGAC GCCAAAGTGG GAAAATTTTA CGGACTTTTT
98341 AGAAATGGTA AACCAATTAA CACTATACGG GTTCTATTTT TACGAGTGTC TTAACCAGTA
98401 CAGCCCCACT AGCATATCGC TGGCCAAAAT TCAAAATATC TTAAACCGAG TCGATGCCGA
98461 ACAGAGCGAT CGCGCCCTGT GGCGCACTCC ATTAATCGGG TCTTTCCCGT TCCCCTGGAA
98521 ATTGAACAAC GTGCTGGCGT TTTTTAAGCC TAGCACCCCC GTGGCAACGC TACAAAAAAT
98581 CTACAAGGCA ATACCGTCGT ACCTAATGAG GTCTTTGTTC GAAATAGCGG CCAACAAATC
98641 GTGGGGTAAC ATCGCGCTCG CAGAGAGCGC CCCGCTAACC GACATACAGA CCGCCGAACC
98701 AGACCAGGGT CCCGTGTCCG CCCAGGTTAT CGCAAAATAT TGCAGCCGTC TCCAAATAAG
98761 CGCGACAGAT TATGACGCAG CCATCGTCTC CAGCCCGGGA TTCGCCGCCG AGTTCATTAA
98821 AACAAAACTG TATCCGATCC TTAGCGAGGT GCTCCGAAAC ACGTCCAAAA AAAATCGATC
98881 GCTATTCCAA ATTAGATGGC TCATAGTCTT CGCGGCAGAG GACGCCAGAG ACCTAGCCCC
98941 TATCAGACGC TCGCTGGCCC TGGCGTATTT TCAAATAATG GACATTTTGG AAGAAAAGCA
99001 TTCTCCGGAG TCGTTTTACA ACCTGCTGGA CTATCTTCAG GAAACATTTA GGTGCATACG
99061 ACAGGTGATA CCGGAAGCCA CGTGCCCACA AGAATTTCTA CAATACTTGT TTACTTTTCA
99121 AAACATTCCA ATAGCAGCGA GCTTCATTCA AACATCCATG ACCTTTGTAG ACGACCTGAA
99181 AAACGGCATC CCCGGTATAC TGGACCTTGT CTCCTTAGGT GCCGCGTTTT ATAACATGAA
99241 ACTACTGTAC GATTCAACGC TAGACACCGT CGAGATTCCA ACGGAAGAAG GGCAACCCAT
99301 CGTCGTGAGC ATGTTCGTAT TCAAATCCAC GATTCGCGTC CTCGAGAAGC TCTTACAGGA
99361 AGCCGTTATC GCGTTAACTC AAACGTCCGA ACCGATGTAC GCAGCGCACA TCCGTCTGAT
99421 GCAACACCTC ACGTACATGC AAAAGATCGC CGGACACGAA ATAATGACCA CACAACTTCC
99481 ATCCGTGTTT CACGAAATAC ACGAGGGATA TTTGCAGTGT TTTAAGCGTT TTAAACGTCT
99541 CATGTTACAC GTTACGGGAA GCTGCTGCTA CTCACTGACG CGATACTTTG GATTCCTATA
99601 TCAACCCCCC CTAATACCCG ATACCATCGT ACAAAAAATT TTAAACTTTA ACGACAAAAC
99661 GGACACAACC GACGACATCT TAAAGAGCCT GTCACAGCCC GTGAGACAAG GACCTCTATC
99721 GGCTGAAAAC GAAAGTAGCA GTCGACTCTC AAAAAACAAC GTTGAGCTGC TTCAAAAACT
99781 GTACGACGAC TTTCGGACCG CCTCCACAAA CAATAACCCC ACCTCTATTA AACTTGAATA
99841 TTCGGGTAAT TATAACGAAA CACAAGTGTC CGTAGATTGG AGCACATATA ACCTGGTGAC
99901 ATACACCGCA CCCGACGATA CGTTAAAATT CACCCCGGTT AACACGGAGG CACTAGATCG
99961 CATGTTTGCA GAATAAACGA TGGAATTGCC GCCAATTTTT TCGAAATTTA AAATAGAGGG
100021 CGTGGCAACG ACGCACCAGG CCGACTGCAG GTTCGGACAA TACGCCGGCT CGCAGTGCTT
100081 AAGCAACTGC GTTATCTACC TCGCACAAAG TTATTTCAAT AGAGAGTCCC CCGTGACGGA
100141 CACTAACGAC CTCGACGACG TTTTACGGCA AGGGGCGACG TTGGACTTTA TACTGAGACG
100201 GTCCGGGACG CTCGGCTATA ACCAATATGC CCAACTGCAC CACATACCCA GCTTTATCAA
100261 GACCAACGAG TGGACCGCGG CCATCTTTCA ATCTCAGGAG TACTTTGGAC TAATTGGACT
100321 TGACGCGGCC ATCCGCGAAC CTTTCATCGA ATCCCTAAAA TCGATCCTAA CGCGAAATTA
100381 CGCCGGCACC GTTCAATACT TTCTGTTTAT ATGCGGCGAC AAAGCCGGGG CCGTAATTAT
100441 CAAAAACAAA ACGTTTTACC TGTTCGATCC CCATTGCGTA CCCCACGTAC CAAATAGTCC
100501 GGCACACGTA ATCAGTTCGT CAGACCCCAC CGCCATACTA GAGTACGTGT CACCGCCAGA
100561 CAGAGAATAC ACTGGCAGTT TTCTATACAT TATGCCCAGC GAATATGTCA ATCCAGAGCA
100621 CTACATCACC AACCACTATA GAACTATAAC GTTCGCCAAA GTACACGGCC CGCACATAGA
100681 TATATCCACC GGCATAGAGC CGTGCACCAT CGAAGACATC CCAAGCCCGC CGCGATCGCC
100741 GGATGTGACG TCAAAATCAT CCAACCTCGC ACGCGTGCCG AGAACCACCA CCGACACGTC
100801 GAGCGCCAAA CCCCCACCGG CGACGCTATC CGGTTTACGG GGCGCGGAAC CACCGACAAG
100861 CTACCCAGAC CCGGCAACCA ACGACGCGGA CACAAAACTC TTAACTCCCG CTCCAGCGCA
100921 AACCGCCGTC GACCACCCCG AATTCCAAAC AACACCTGGA GCCACGCTAC TTCTTTCAGA
100981 ACTGTCGGCA TCCAGGGGTC GGAAACGCAA GCTTTCCAGC CTTCAGCGAT ATTCGGATTC
101041 CGACGAAGCG TCGTCTGACG ACGAAGGGGC CCCACGTAGG CGCGTACACG ACGATGCGAT
101101 ATCCGCCGAG GTGATCTGGA TGGACGACGA TATATCTCCT CTATATTCAC CGTCGGCGAC
101161 TCCATCGTTT GACGACGTGT TCGACAGCCC CCCAATGAGC CCGGAGTTTA CATACGAAGA

101221 CGCGACAGAG GACACGGATG GCGCGTTTCT AGAACAGATC GCTCGAGACG CGGAAACGCC
101281 GTTCTCTGCG TTCGACGACC TTATAACGGA CCACGACTTT TCTTCCCTTG ATAAAAAAAT
101341 AGAACAGTTA ATCAAGTACG AAGCGCCCTC GCAGCACCTG CCAAACATCT CGGACAAACA
101401 AAACGGGCGA GCCGTCCGAG AAGCGGCGGC CCTCCAGGCG ATGGACAAAA TTATGATTAA
101461 TATCATACTT GAACACGGTC TAATTACAGA CGCGCAGGCC CGGGGACCGT CCGCGTGCAA
101521 AAACGTTCTT CAATTTTTCA TCCTGTGGGG AGAAAAACTC AACATACCAA TAAGCGACGC
101581 CAAGCAGGTC CTAGAACTCG ATCTGCAACT GATACCGTTA CATACGGCTA TCAGCGAAGG
101641 AAAATTCAAA CAGGGGGCGT TCAAAAAACA CCTAACAACT AAAATCAACC GGTGTCTGGC
101701 GTCAATGAGG GCCACGCACG CAGATGCGCA AAAAAAACTG GCGTCGGCTT TCAACGTCGA
101761 AGGCTCACAG ATTTCGTCCA GCGAAGCGAA AATATCAGTC CGGGCGCTGA AGGAACAGAT
101821 CGCCAACCAC CTGAGTCCAG GCTTTTTAGC GGTCTACTCC GCGGACGAGG TAAAACATCT
101881 ACGGGATAAA ATTCAGGACC TAAAAACAGG CATCGAGCAG CGCAACAAAG AAATCCAACA
101941 GGAAGAACTG TTTTTTGATG CCATGCTCAC AGCCCTGGAC ACGTTCCAAC CCCCTCCGAA
102001 AACGGCATTT CCAATGGAGA TCTTTCCGCA CCGTAAAACC GAAGTTATGC TCGACCACCT
102061 GGCGTCCATA ACCACCAGGT TAACCGAGGA CGCCACCGAA GCCCTCAACA ATTACCTGGA
102121 GACCCCGCCC GACCAAGGAA CGCACATTAC CAACATTCCA AACTTTTCAT CCATCGTGGC
102181 AAATATCATA TCCACGTTAA AATCCTAAC GTACGCAGAA AACGACATGC AATTAAACGT
102241 AACGCCCATG GCAACGTACA GGCGTCAGCT GTTGTACCTC GGAGGCGAGT TAGCAACCAT
102301 CTTTAATTTA GAGTGGCCAT ACGAAACCGT GCCACCGGTT CAAGAACTGC CCCTCGTGGC
102361 GCGGGCGAAA GCAAAAATGG AATCGGTAAC AAAAATGGAA AAGAACCAAC AGGCTCTCGA
102421 CCAAATACTG GGAGACGCCG AAACGTTACT TGACACAATA ACCGCAACAT CCGGAGATGA
102481 GAACCCGGTC CGCGCCATGT CCATACCGAT ACTGGAGACC TACATTACAA ACGCAGGCGC
102541 CCTGATAGGC AGTTCTCGAA ACCAGCGGTT CGAAAAACTC AAGGCCGCCA TCCACGACCT
102601 GGCATCATCG GAGTCGTTCA TAATAATGCT GCTAAACAAC ACGCGGCTCG ATAACATATC
102661 AGACAATCTG GCCAAGATCG ACGGCATCCT GACCAACAAC ACACGTTTTC TTTCAAACGC
102721 CACTGTTAGC AAAACGCTCC AAACGCTGGG AGGCAGCCTA ATACGCGAAT GCGTAGAAGC
102781 GCTAAATAAA AGGAGCCCCT CTTCCCTCAA CAACGCGCGT CTCCTCGCGG TTCAAACCAT
102841 ACTGGGGCAC GCGTCCGTTC CAGATCACGA GACGCTGACG CGAATCGTTT CCGGCGTCGC
102901 CAGCGCACAA AAGGAATCCG CTGGCGATGA TCCAGATAGG TGGACGCGAG TAACCGGTCA
102961 CCTAAACGAG CTGAAGCTCG TAACTACCCA ATCGCGTGTC GACAAAGCCA CCAGGCGCAA
103021 ACTGTTAATG ATAATAACCC GTGACCTCAA GGAGGCGGAG GTGTCTCAGG AAACGGTCCT
103081 GGAAACACGG TGGCAAGAAA ACGTGCTAAA GTTTCAACCG TCGACGTCCA AAGAAATCGA
103141 AGACTTTTTA CAGTCGGCAC CGTCAGCAAA GGCCCGAAAA TTCGCAGAAA AACACCTACG
103201 GACGCTGATC ACCCAATTCA ACGGCCACGA GCGACCGCCG TCCGAGGCCA CCGCCGTTCC
103261 CATGGACTAC ACGCCGACGC CCATACCCAC GCCACAGGCC GTTTCTACGG CTACCGCGGA
103321 AAAGGGAAAG GCCGCATGGA ATAAAATTCA ACAGGCCTTT CAGGATTTCA ACTTTCACCT
103381 CATCGACGCT TCGGATTGGC AAGAGATGGC ATCAGAATAC TCCAGACACG GCTCGTCCCT
103441 TCCTGGTACG GTTGGACCAA AGCTGGTGCG CTTCATGGAG AGCATCTCAA ACACCCTGGA
103501 CGACATCCTC ACGCAGAAGC TGGCATCTCT GCTTCCAAAC GGGCCCGCGT TCAGACCCCC
103561 AGCGTTTGAC TGGATCGCGC CCTATCAAAC ACGCGTAAAC GCGTTTCTAA AAACCATAGG
103621 CCTGCCCATG GTGCGCAACC TGGCGGACAA GATCCATCAC CAATGCCAAA CTGTCAGTCA
103681 CGCGGTGCAA TCCGCAGACC TTCAACAGGC CACGGTGGGA ACAAGTTTAG AACGACCCGC
103741 GGCCGAATAC TGTCGAATAC TCTCTGACAT GCAAGTCGCG TTCAACGACC ACGGAATCGC
103801 CGTAAGATCG GAGGCCGCGG CGTACACGGA CGCAATCAAC TCGCCGGCCA ACGTCGTGAC
103861 TCCCCCGAAA CCCAACCTAG AAGCCCCCAA GAAGCTAATA ACGGCAACTG ACGCCCTAAC
103921 CGTCGAGGAC TTTCCAGATT TCCTAAAAAC GTCAATCCTT CAACAGGAGC AGCGACTCAT
103981 TGCGCTCCAG AGAGCGGAAT TTCAGCAACT AGAGGCCAGC ATCTCGGCGG CCGAACGGCT
104041 CCGCCAATCC ACCCGTGACG AGATCGCAGG CAAGATGGCA ACCGCTATCA CGCAACTCTT
104101 ACCCCGCGCC CCCGTCGCAA TATCATCGAG ACCGTTGAAC TTATCAAAAC CTATAGACTT
104161 TTTGAGTTCA ACGGTATACG ACAAAATCCT GGACAAGGAG CCTTACGAGA CAGCCATAGC
104221 GGGATTCGCG TGGCTGGAAA TCGCGACAAA ATCCGTAATG GTCTACAGTC AACAAAACGA
104281 AACGCAACAG TTAAACGTAC TGCTGAGCGA GGTAGAAAAA CAGAGCACCG TCGCGCAGCG
104341 TCTACACGAT TTGGAACTGT CGGCGAAAAA CACGGACGAC GTAAAGGTGC TGAAGCAGGC
104401 GCTAGACGAA CTCGCGCCCC TCAGGGTAAA GGGCGGAAAA ACCACCGTAG ACGCGTGGAA
104461 ACAAAAACTG GAAAGCATAG AATCCCTGCT TCGCGCCACG AGGACGGCAG GCGAAATATC
104521 ATCGGAGCTT GAACGCATCG GCACACAGGC GGTTGGCACC ATCACCGTCC GCGATTTAGG
104581 AACGCTCTCC GATCAATGCC GGGAAGCCGC AAATTTCCTC AGACAGGCCA GTCTACCCGA

104641 AGGCTTCTCG GACATAGGCA CAAAACTCAG CGAGCTTCAG GCGTACATTA AGTACAAAAA
104701 ACAGTTTCTG GAGCATTTTG AAACAACCCA GCCTAACGTC TTTCAACGCT TCCCGCTATC
104761 CCAAAACATA ACCGAAAACG TTCCGGCGCG CCCGGCGATG GACTCGGTGG CCAGACTGAC
104821 CAATCACCTT CACGTGCGCG GCAGCGCGCC CCACTTTACA ACGTGGATAG AAACGCTACC
104881 GACCGTCGAT CCGGAAAAAC CAACTCACGT CCCGGCGCAC GGAGGAGCCC CTCTGCACCG
104941 CCAGATCACG TACTCAAACG TCCTAGAGGC GTTGTTTTCA TTATGTTCCA CCACGCTAAC
105001 CCCGGTTCCG ACGGCCCCCG GTCTGGAAAT CGCAACCAGG GCACGCCGCG GGGCAGAGGC
105061 CGCAACGTGG ATGGACAGAC AGTGGCCCGA CATCGCTCAG ACGCTCCAAG ACGTTCTCGA
105121 CACGTACGAA CACACCACCG CCCACGCAAA CCGGGACGCG GCATTTAACA CATTCTTGGC
105181 GATGTGCGTT TTTACGCAAA TCATCAGGGG CGCTAGCAGA GCCGTGACGC TCCCGAAGTT
105241 ACCGAGCACC GCCGTCGATT TTCCAGAAGA GATCGTTCTA ACACCCAGGG AATGCACAAC
105301 ACTGGTCACC GCCATGTGGC CCACCCTGGC GGCCGCAATC TTACGATTAA AATCCTACTC
105361 GGAAGCCCTA GGACTAATGA GTCGTTTCCT CCCGCTAATG TTCCAGGCGC TGCCGCACCT
105421 AACGCTAGAG GCCCAGGTCA AAAACGGCCC ACATAACACC CCGCCTCAGT TGAGATGCTT
105481 TGCCAAAACA GAGGCAATTC CGTATTTCCC GGCGCAATGG CAGTCAGCGA ACCTAGAGCA
105541 GAGCCTGTGG GGACAGACGG ACTTTTTGCA AATCTGCGAT AACAATCAAC GCAAGGCCAG
105601 GGTGGCGGCC GTCACCTGGG CGCTCACGAC GATAGACGGC GTGGTTTTGG ACCAACTGTG
105661 GTCCACATTT AAACCCATGA CAGCCGCGTC AGACGACACG TACGTCGACC TAGTCGAGAC
105721 CCTACACCTG ACCACCTTTG GCCCGCGCGG TCCAACGCCG AGGCGAGAAA CGACCACCGA
105781 GCACCCGCCG TACGAGTACG GACAGCCCAC GGGCTACTGC ATCTCGGGTC AATCGACGAC
105841 GCCGGTCCAG GCTTCAAACA CACCGGTATC CGCTTTCGAG GCGGTGCTCG GAGCAATGGT
105901 GTTTCACGTA CCGATCAGAA TATTTTTGGC GGCCACGCCC AAGCGCCTTG GCCAGGCGCG
105961 CGGCGGCATG GGGCTCCTCA CACCCATCCT GGAATGCGTC CCCGACGTCG AGCCCTTCAA
106021 AAGCCTGTAT AACGCACCCC GCAAACCCGT GCCCATTGAA ACGCTACCCG CATCCCTCCA
106081 CCCGCACGAC GAGCGACAGG TCTTTCTGAG ACAGGCACAG TGGCTATCCT ACCGATTCAC
106141 ACCACACGAA GCCGCCCGGT CGTCGACTCC GCCGCTTCTG GTGGTCATAG ACCCTGAAAA
106201 CCTCGTAACG GCAACGTACT CCAGTGGCGG GCCTGCAAAT TTCGAGAGCA GGCCGTTTTA
106261 CGTGATGCCC GGACCATACC CCCCAGACTG GCCAAAAACG CTGTCGGTAA CATCAAACAC
106321 GTCCGTGACG CACCTCAGCC ACGACGAGAT ATGTAACCTC TTTACTACGC TATCCCGAGA
106381 ACACGGGACC GTGCAAGGCA GGGATATCTT CGCAGCGGCT CCGACAAACG TCACACCGGA
106441 ACAAACCGCC AATCCTCCGG CATGGGAAAC GGATAACCGA TTAATAACGC AAACAGAAAC
106501 CGCCAAAAAA CCTCATATAA TTCCTGCGTC TCCTAAAGCG CGGACAGATC CACCGGTGGA
106561 AACCACGACC CACCATTCAC AAGGGCAAGC GTCGCAACAC GCAAACAGCA ACGTAAACCA
106621 GCCCGGTCAA ATTACTTCAC ACGCGTCACG TAACACACCG TCAACCGCAC CTCAGGCCTC
106681 ATCTTCACCG GAAAAATTCA ACACGCAAAC GGTGCCTCGA CTAATATCTC AAACGTCGGA
106741 AACGGCCCAT ATAAACCAGC CAGCCTCCGG CCAGGTCACC GAACCAAAGG GAATCTTTGG
106801 GACGTATAAA CCCCGAGTGC TCACCGAACC CGCCAAACCC GCAAACGCCG GCGTAGCCTC
106861 TCGCCAACCA GAGGCAACCA CCACGGTCCC CAAGTTACCG ATTAATCCAC CCACCGCTAG
106921 GGTCTTTATA GGGACCGCGT CCAAACTCTC GCCAGCCGTC GAAGAGAGCC ACGGCGCCAC
106981 ACCCGACGCA CATCAGTCGA AGATAGATCG GGAAAAATAC GCCGAGAGTC GGCCTCGCCG
107041 CACCCCACAC CTCGAAGAGG GGCCACGGGA GCCTCACGTC AACACTCCAA CCAGCGCACA
107101 CATAAACGTC CCCTCTAGCC AAGGTCAAAA AACAGTACAC GGGCGCGAAA ATCCCGGCCT
107161 TCAAACAGCA ACTCCCAGCG CCCCCCAACC AACCGCATCA AACCCGCGCA TTCAATACAC
107221 GCTCCCCAGA ACGGACGGCC GGTTGCTTCA CGACGAATCG GAGGTGGAAT CGACCCCAAC
107281 CGAGGAGGTA AAACGATCGC CAAAAACACA AGATGTGTCT CACGGGCCCG AACCGGACGA
107341 CTCCAGGTGG ACCGCCCCGC TCGGTCCAAC CATAGAGATT CATCGACTGG AACACCCCCA
107401 AATTCTCAAA AATATAACAT CACTCACCGT CCCCACTCCC AGAGTCACCC CAATCCCTCC
107461 CACTAACATC TGGATACCCC TATCCCACGT CAACATCCAA CACGAAGAAA TCACACGAGC
107521 CAAGAATGTG TTAATGCGAT TTATTCAAAA CGTACGAAGA AAACTTCAAG CGTCGTCTGA
107581 CGCTCTATCC GAGGCTATTG CCAGAATAAA GTTTTTATAT CTGTAACGCG CCCATCTCAC
107641 TTGCTTTTTT TATTTTGAGA CGAGCGTCTT GTGTCCAGAG TAGTTGCGTC GCTAGGTGAT
107701 AACGAAGTGG ACCCTATGCC AGAAGACGCC ACGCTCCCGG GTCCGCCGCC ACCGGGCGCG
107761 GGGCCCATCG AGCCTATCAA TGAATGGGGC CCGCTGGAGA TCGTAGTAAA GCTATTTGAC
107821 CCGAGGGTGG AGGCCACCGG GGCAACGCTC GGGGCGGCCG AGCCGGACCC GGACAAAACG
107881 CCCGAGATAC TAGAACTAGC GTCCTTTTTG CCGCGAAGGC CCCGGAGGTG GTCTTTGCGC
107941 AGAATTCCAT TCTTCTTTTG CATATACATG TCGTAATGAT GTTTGGCCGT TAAAAACACC
108001 AGATAATTAC GTTTCGCGAT GGCATACTGG GCGGGAGACA TGTCACCCTG GGGAAGGTTG

108061 TTCATCTCGG CAACCAGCGG GTGATTTGGG TAATCGTGCT CAAGGCGTCC CTGAACGATT
108121 GGCTCCTTAA CCCGCAAGGA CGACATCTCT TTTAACTTCT ATTACACTTT CCACAGGACA
108181 GGGACGATAT AGACGAGGTC AAATAAAACA GCTCGGCCAC ACGCAAATGC TTTAATAGGC
108241 CGGTCGCGCA GTCCGGCGAC GCCAAACAGG GCACGACGAC GCTAACCAGG GAGTTCGCGT
108301 CTCGTATGGC ATGTGCCGCG TTGTTGGCGA GCACGCACCG TAAATAGGGA TCTCCAACAC
108361 ACGTGATCTC GAATAGAGAT ATAACCCGCA TGTGCGATCC GCCACAATAA GAGCAATATA
108421 CGCGCCCGGT GGTAGCACAG ATCGAAAGCT GCTTCTCTTT TTGGTCGCGA CTGAAAAACA
108481 CGTTGGTGGG TGGGAAATTT ACGGTTTCAA ATTTACCCCG TCCGAAATTC AAACAGTAAC
108541 CGCACTCGAG GCACACCACC ACCCTCGGAG CTGGCACGGT CTTCTCCAGT ACGCTCCTGG
108601 CCACCACCTG GGACCAAACA GGTAGAGAGA TACACGGAAA CAGTACGTTA TACGCCAATA
108661 CTTTTTGACC CAGGTCGCGG GATATCTCCG TCTCGGTCGA CTCCCCTATG GGCAACACAA
108721 CACGGGACAT GCTCAGCAGG GCCCTAAACG TCAGGCTCCT CAGAAGGGCG TTAAACGGGT
108781 TGCCGCACGG GACGGTCGGC GCCAGTTCTC GCAGCGAGGC CAGAAGTCCC GCGTCCGAAG
108841 GGCCCGGGAC ACTCTCATTC AGGTTAGCTC CCAGACGTCT GGAAATGGAC GGAACGTTCA
108901 ACTGCATCGA GACACAACCG CCCCCGTTCC ATTTCTTCCG CAAAACGGGC AGATCCAACG
108961 CGTGCTGTGG CAAACAGGTA ACCAGGGGAA ACCGCTGGCG ACAGTTAAGG GTTTTGCACA
109021 CGAGACAACA CGCCCTCTGA AACGACACAA CAAGATACCT GGACCACGCG CTCCTGGGAA
109081 CGGCCGTTAA ACTTAAACCT TCGTCGCTGG GACAGCCCAC GCCGGTGCAG ATACACCTCA
109141 GCACCCACGC GTACGCCTCT AACAACGACC GGCCGATATC GTGCAGCCGC GATCTCACGT
109201 CGCCGTTTTT AGGTGGGTTA TCCGGCCTCC AGCCGGTAGC AATCTCGTTC AGGGCGGTCT
109261 GAAAGGATGG GGCAGAAATT AACGTGCAAG CCCATTTTGG GGGTCGTCCG TCCCAGGCAC
109321 CGAGCCCGTA CGTCACAAAA CACACGTAGC ATTCGGGGCA TAGTCCGATT GAGCGTATAG
109381 AGGCCGCGAG ATCTAAGCCC AGCCGAGAAC CATCCAGCCA ACGATGGGCA GGATAAGAAC
109441 CGCGCCCATG ACAGGCAGCT TCGTCTTCAG CCAGGTTAGG CAAGCGGCCC GAGGCCATCC
109501 CCCAATTGTA CCGATTGAAT TGGTTAATTG GTCATCGCCG CACGCTCTCC GCGCCCAATA
109561 TCCTTCAACT CCGACCCCGA AGGGCGGGCC ACGGAGCCCG TTCGCCTCAA CGCACCGAGG
109621 CCGTCGGAGT GACCCGTCAG GGCAAAAACC GTTCTAAGAA GGGTTTTTAA CCGTTTAGCG
109681 CTCTTTGGAG TCACGACCAA AAACTGTAAA ACCTGTCGGT GCTCCGTAAA GTAGGTGCGG
109741 CATATGACCA TGGAGCTGTA AACGTTTAGG TCTCCGGAGA AAACCAGACG TGCCTTAAAT
109801 TTCATAAAAT CGTCCTGGCC CAGGGACACG GACGAGTTCC TCTCAAGATA CACGTCCGAA
109861 TTTATAGGCA GGTTTTTTCC AAACTGGGCA TCGGCGTCAC GTGGCTTACA CAAAAAACAT
109921 TTCAGCGTGG TGGCCAAACC GTTGTTGATA ATTACAAAAC ACGGGGCAAA CGGGTAGGCC
109981 AGTCTCTCTA GTTTGTGGAG CCAAAACTTA TACACAAACT CGAGATGATA GACGCAGCCG
110041 TGCTGCAGGC GCACGGTGCA CACGGGGACC GCCCCGCCTT TAGCGTATAC GGGAGCCCCG
110101 TCCTGACACC TCTCCAAGTC CAGGGAGATT CCAGAGGGTC CCAGGTAAGA GACAACTAAA
110161 TCGCACAGCT CGTCAACTAA ACGTTTTCCG GAACTCATCG TTATAAAGAT CCTTTAGGTG
110221 CTGTGCGTGG CTCCCGTAAA AACCGCGTCC GTGCTAACGA TTTTGTGAAT GACCTGTTTT
110281 ACGGCGTTTA CCTTGGCGTC CAGGACCATG CAGTGCTCAC AGTGAGCTGA CCGCGTCTGA
110341 GCACGATGAC AGAGGAAAGT TTTTAAATAC TGACAGTAGT TAATGGCGTT GAGCCTGGAA
110401 TATATGGTGG GAAACATAAT TTTCATGTCA TCGGGCAGCA GGGACTCGAA CGCCAATAAA
110461 TCGTCACCGA ACATCACGTG AGACAGAGGT AAAAGATGCT CACCGCCGGT ACCGCGTAAC
110521 ACGCGACCAG CCACCCCTTC AAATATTTTA GCCTTAAAAA GCGGGCCCCC TAAAGTCGTC
110581 CAACTCAGCT TAAAAACTCC TACCCATTAT GTTCGTCCCG TGGCAACTGG AGACCCTCAT
110641 GAGACACTGG CCCTCTCTGC GGGGACTCGT AGAACAATCC TTCCTCCCCG GTACCCCGGA
110701 CGGAGCTTTT AACAGCCCGG TATTAATCCA CACTCAGGAC TCTCTACAAC CCGCCTCATC
110761 GTGCAGGGTG TGTAGCCTCC TGTTCACTCT GGTCCGGACA TTCCCACCCC CCGACTCTTT
110821 CTTCGAAGAC TACGGCTGGT TGTGCCTCAC CTGCCTATAC GCCCCCCGAT CATGGACGGC
110881 TACCCTCATG GTGGCTGCCG ACCTTTTGGA ACTAACGCAC GTGTACTTCC CGCAATGCGT
110941 GAAAGATGGG CCAGTATACA CCGCCCAAAG CATCCTCGGA ATCGACGTCC AGCTGCACTT
111001 CTTCGCAACC CGCTGCTTCC GACCCATCGA CAGAGAACAA ATACTCCACA CATCTCATTT
111061 AAATTTTTTA CAAACCGAGT TTATTAGGGG CATGTTAGAA GGCACGATTC CGGGATCGTT
111121 CTGTTTTAAA ACGTCCTGGC CGCGCACAGA AAAGGACGAC CAACAACCTA CCGTTGCGTG
111181 TTGTTCCGTT GGCCGCGGAA GTCACACCAA CCGGGATAAC CGCCTACCCG AGGACCTGGA
111241 AGAGGCGTTC AACTCCACGA ACGCCGAGGA AAAGCCCAGC CTCCTCGGCG TCTTTTCGGC
111301 AACGTGGGCA GAATCCCAGC TTCTTGGCTC CGACACACAA CAGGCAGATA CCCATTTACA
111361 ACCCTCCGCC TTCCCAACCC CAGAAGATGC TGACCAATCA CAGGGCCCCT GCCTGATGCA
111421 CCCAACGCTC AACCTAAAAA CAAAAAACCA CACCGCATCC ATATGCGTTC TATGCGAGTG

111481 TCTGGCCGCA CACCCGGACG CCGGTCCGGT TCTGAAAGAT CTGCGTCGCG ACATTCTGGA
111541 AAACATGGAA AACAACGTTA AGCTCGTCAA TCGCATATCG TACATCCTAA ACGATCCGGA
111601 CTCACTGTCA CACGTGCGCG ACGAACATCT GCGCGGCCTA ATTAAACGGT GCTCGGCACA
111661 AGAAATCCAC AAGCATTTTT TTTGCGACCC GGTGTGCGTC CTGAACACGT ACTCGCACTG
111721 TCCCGCGGTT TTATTTAAAT GCCCACCTCC CGAAAAGTAT AAGAAGCTCA AAGCTCGTCT
111781 GGCAACCGGA GAGTTCCTAG ACTGCAACAG AATATTTGAC TGCGAGACCT TACAGACCCT
111841 GGCCGTCCTC TTTAAGGGGT CTCAACTGGC CAAAATCGGC AAAACCACGT CGCTCGAGAT
111901 AATCCGTGAA CTCGGATTTC AACTGCGTCG ACACAACATT CAAATCACCC ACCCGTTTCA
111961 AACCTCCAAC CTATACATTT AATCTTCAGA AGCGCACCAG ACAATGCCAA AACAGCCCAG
112021 AAGTCGATTG GCGTCTCGAG CGCCGTACGC ACCTAGCGTC AGGCGACCGG ACGGGCCCCA
112081 GTCCACGCGA CCGGCATCCA GGCACGGCAG CTGCAAAAGC GAAATCATGC AGTGGAAAAA
112141 GTTAGTTTCA GACACGCAGT TTTTTTCTGC CCTAACGCGC CGCCACGAGC TGGGGGTGGA
112201 CTTTTTAAGA GAAATGGGGA CCCCGATATG CACCTCAAAG TCCGTTATGT TGCCGTTAAA
112261 CCTAAAAACC ATCGCCCCGG GTCGGTGCGT CTCTCTCTCA TCATTCGGAC ACTCGTCAAA
112321 CATGGGGTTC AACTGTTCGT CGTGCACGCC AACTGACAGG TCAGCGGTGT CTCTGGACGC
112381 AAACGCGCTC GGCGAAGATT CCGCCAGGAA AAACAGCGAG CTGTGTTCAG TGGCGTTAAC
112441 CTTTTACCAC CACGCCGAAA AGGTCGTGCA GCACAAGGGC TTTTACCTGT CTCTGCTCAG
112501 CCACTCCATG GAAGTCGTTA GGAAAAGCTT CACGCAACCC GGGTTGCTCT ACGCCCACCT
112561 AGTCCTAAAA ACCTTTGGCC ACGATCCTTT ACCTATTTTT ACAGTCGATG CCGATGAGAG
112621 ACTCGCACTC TGGGCGGTGT TCCACACTAG AGACCTACAC CTGGGGGAAA CCAGTCTGCG
112681 ACTCATTATG GACAACCTTC CAAATTATGA CATAACGGTG GACTGCATCA AGCAAACGTA
112741 CATAATGAAG TTTACACCCT CGCGACCGGA CAACGCAACC GTGACGGTTC CTGTCAACAG
112801 CATTTGCGAG GCCGTGGCCA CCCTAGACTG CACCGACGAG TTTCGAGAAG AAATTCAAAG
112861 GGGCACGGCC ATCATAAACT CCCAGGGGCT ATTGTAACCT TTCCCTAGAC GGAAACAGAT
112921 GTAATTCACC TATTCAAATG TTAAGTTTTT GTTTTGAACA TATCACCAAT AAAAACAACT
112981 TATGTTTACA CAAATTGAAG AGCGTTTCAA TTTTACCATA AACATAGCAA AAACCACGGT
113041 AACTAAAACT CAATAGCGAT ATTCGGATGG GATCCCTAGA GAGCAATCGA CGTGCGCCGG
113101 TATTAATTGA CGGGAACGCA CGTTAACTGC TCTCCACGAA TCGCAAACTC CGCGTTTTTA
113161 GGACCATTCG ACGCCGTTAA TAATTGGCAG TAAACTGCGA AAGGGCGTAC ATCGCAGTAG
113221 TGATTTTACG ATGTGCACAC CTTTCCGCGG TTCACCGCCA ACAACAAAAG CGGGTCACCA
113281 GGCACCTAAA TCGCCCGGGC CTTTTATGCC AAAGTACAAA AGGGACCGGC GTTCTACGTT
113341 CACGAACAAC ACTGTTGATC GCAAATACAA TAGTGTTGTT GGTGACCGCA AACTGCCAAA
113401 TCCTAAAATG CATGGACATA GTCCTCAAAA TCTAGAACGA ACTTCCAAAA ACATGCGGCA
113461 ATAAACAACC CCTTGCAGTT AATGCACATC GTAACGAGAC GGAATCCAAA GAACACGCCC
113521 GATCTTAAGA CCACGGGCAC GTGTCTTTGG GTTCCGGGGG CGTTACGATG GAAGTTCATC
113581 ACCAAGTATC CCATCAACCT GTTACAAAAG CGTAAAGATA ATCGGTTTAC AATAAACAAT
113641 AAAAAATAGT GGACGAAGTT ACATTCCGAC ATACAAGCGG ACCCAAAAAA ACACCCCAAC
113701 GCAAAACAAC TGGAGACTGT GTCTTTGGGT CCCTTCCACG TCGCAATGAG ATTTCTGGTG
113761 ACCGTGTGCT GCTTAACGCG TTACAGTTGT GTTTTATGGC TACGTAAACA CACCAATCTA
113821 GAATGCTAAC ACGTGCACCG CGCTTAGGCG CGTTTGTGCA ACTAACTGCT GTCAGGTTAC
113881 TACTCTTATT TTTAACAAGT AGCATACCAA CTGCAATTAA TTGCACTAAA CCCAGTTTAG
113941 TCTTTTTAGT CGAGTCGGCC CATAGTCTTC ATGGAGAACC ATCATTCATC TCCGTCAACG
114001 GCGTCGGTGT TACCGTTGTT ATTGCGGGCA AAACAGGCTC GGGGGAAAAA AGAGCAACAT
114061 ATCGTCAGTT TAAAACAACT CAAGTTCTAA AATCTCACCC ACTCGTAAAC AAAAAACATC
114121 GCAGTTAAGG TGTATGTACC GAACATAACA ACACAAGTTT TTTAAAACAC AGCTGCGGTA
114181 AGTAACCCCC ATTGCCACGC GTGTCGCGTG CTAAGTGTTT TTAAAATTAC ATTGTGCGTT
114241 TTACACACCG AGCAGTAATC TCAGGAGGGC GGTTAACGAG CGATATACAT ATTCCCTAAA
114301 CACGGGAACG CGCGCTGACC GCCCTCCCCA AATCACAACA CGGGACTACA AAGCCTAGTG
114361 TTAATATAAT CAAATTAAAA AACCACAGAA ACCTTTAGTC GTGCGCAAAC ACTAGCAAAG
114421 GTACCTAGAG CTTTCCCCTA TACTTCAAAA AACAGCGGTG GGTTATTTGA CCACACGTTA
114481 AGTAAACACC CGTAAGAATT ATTCCCGTTT ATCAAAATGG AAAAATTAAG GCTTTGCGTT
114541 AAAATCTGCT AACGCAAAGG GCACTTAATT TTTCCAGTTT GGACTCGGAA CTTTACCGCT
114601 AACGTTAAAA TTTAATAATG CAAAAGGCAC ATAATTTTTT TTGTGTAACC TCAAAACTCT
114661 ACAGCTAAAG TTAATAACGG CGACTTTGGG CAGCGAATCA GTGTCTGTCG CAAACCCTGT
114721 TAAATTTAAA CACAACAGGG AGCGCGACTA AACACGGACT AACTGCTCAT CGGGCCCTGT
114781 TGAAGAAGTT GGTGTTGCAA TGCATTAGGG AATCTAAAAA CGAAATCCTG CTCGGTAACA
114841 CGGAAATTGT CTTTAGTTCC CTACAGCATC ACAACAATTG TAAACCATAA ACGTACGCGC

114901 GTGGGGTTGG TGTGGTGTGA AATGCTTGGT GCAACTGTTA CATGGGCGGA TTGTAAATGT
114961 GGTGCTTGGT GCAACGGTGG TGGGGTGCAA GTCCCCCGGT GGGGTGCAAG TCCCCCGGTG
115021 GGGTGCAGGT CCCCCGGTGG GGTGCAGGTC CCCCGGTGGG GTGCAAGTCC CCCGGTGGGG
115081 TGCAAGTCCC CCGGTGGGGT GCAAGTCCCC CGGTGGGGTG CAAGTCCCCC GGTGGGGTGC
115141 AAGTCCCCCG GTGGGGTGCA AGTCCCCCGG TGGGGTGCAA GTCCCCCGGT GGGGTGCAAG
115201 TCCCCCGGTG GGGTGCACGT CCCCCGGTGG GGTGCAAGTC CCCCGGTGGG GTGCAAGTCC
115261 CCCGGTGGGG TGCAAGTCCC CCGGTGGGGT GCAAGTCCCC CGGTGGGGTT CAAGTCCCCC
115321 GGTGGGGTGC AAGTCCCCCG GTGGGGTGCA AGTCCCCCGG TGGGGTGCAG GTCCCCCGGT
115381 GGGAGCGGCT CGGCTCCGGG GTGGCTCCGG GTGGGGGCGG CTCGGCTCCG GGGTGGCTCC
115441 GGGTGGGGGC GGCTCGGCTC CGGGGTGGCT CCGGGTGGGG GCGGCTCGGC TCCGGGGTGG
115501 CTCCGGGTGG GGGCGGCTCG GCTCCGGGGT GGCTCCGGGT GGGGGCGGCT CGGCTCCGGG
115561 GTGGCTCCGG GTGGGGGCGG CTCGGCTCCG GGGTGGCTCC GGGTGGGGGC GGCTCGGCTC
115621 CGGGGTGGCT CCGGGTGGGG GCGGCTCGGC TCCGGGGTGG CTCCGGGTGG GGGCGGCTCG
115681 GCTCCGGGGT GGCTCCGGGT GGGGGCGGCT CGGCTCCGGG GTGGCTCCGG GTGGGGGCGG
115741 CTCGGCTCCG GGGTGGCTCC GGGTGGGGGC GGCTCGGCTC CGGGGTGGCT CCGGGTGGGG
115801 GCGGCTCGGC TCCGGGGTGG CTCCGGGTGG GGGCGGCTCG GCTCCGGGGT GGCTCCGGGT
115861 GGGGGCGGCT CGGCTCCGGG GTGGCTCCGG GTGGGGGCGG CTCGGCTCCG GGGTGGCTCC
115921 GGGTGGGGGC GGCTCGGCTC CGGGGTGGCT CCGGGTGGGG GCGGCTCGGC TCCGGGGTGG
115981 CTCCGGGTGG GGGCGGCTCG GCTCCGGGGT GGCTCCGGGT GGGGGCGGCT CGGCTCCGGG
116041 GTGGCTCCGG GTGGGGGCGG CTCGGCTCCG GGGTGGCTCC GGGTGGGGGC GGCTCGGCTC
116101 CGGGGTGGCT CCGGGTGGGG GCGGCTCGGC TCCGGGGTGG CTCCGGGTGG GGGCGGCTCG
116161 GCTCCGGGGT GGCTCCGGGT GGGGGCGGCT CGGCTCCGGG GTGGCTCCGG GTGGGGGCGG
116221 CTCGGCTCCG GGGTGGCTCC GGGTGGGGGC GGCTCGGCTC CGGGGTGGCT CCGGGTGGGG
116281 GCGGCTCGGC TCCGGGGTGG CTCCGGGTGG GGGCGGCTCG GCTCCGGGGT GGCTCCGGGT
116341 GGGGGCGGCT CGGCTCCGGG GTGGCTCCGG GTGGGGGCGG CTCGGCTCCG GGGTGGCTCC
116401 GGGTGGGGGC GGCCTAAAAT CCTTACCGGT AAATTTAGCA GTAAATCCAA CGCAGTAAAT
116461 CCGCAAGCTA GCCGCACAGA GGTGCGACTG CCTGCCAAGG CTCCTGGCGC CTCTTTTATA
116521 GCGCTAAATG CCCTCCCCAA ATGGTTACTA TGGTTTAGTT ATGCAACCAT AATACCAATA
116581 AATGAATCAC CAAGAATACG GCCAACGCGC CAGCGTCCAA CCGCCCTCCA ACCGTGCAGC
116641 GCCGTATATT CGAATTGCCT TCCGCACGAA TACGGGGCGC AGCATGGTCG CGGGAGGCTG
116701 GCATTGCGCA ATATTCCCAG CATTCCAACG GGCACGCAAG GGTCCAGGCT ACCGGGGATT
116761 CCCCAAACAT CATCCAGCGC ATAGAGGTTG TGGCTGAGCT GCCTATCTGG GGATTCCCCA
116821 AACCCGAGAC CCACTGAGAT GCTATTCAGT GGGACTGCAG GCTGGGTCCT CCCAGGACAG
116881 GTGCCTGCGT GTGTGGTAAG GGAGTTCCCC TAATTAAAAT ATTATATTAA TAAATAAACC
116941 AGGCTAAGCT GTAATTAATG AGGCACAGGC TGCCGACCTA GGTACCTGGG GAACCCCCTA
117001 ACGGGATGTA ATTAATTCAA ACCTATATAA ATTCCACCCT GTTAGGGGGA TCCCCACGTG
117061 TACCTGAATT AATGCGGGAT ACAGGCTGTA GGTTTGGTAA TGGGTACCAG GCTGGAGTGA
117121 CTCATACCAT GGGATTAAAT ACCAGGGATA CAAGCTGTAG GTTCGGTAAC GGGTACCAGG
117181 CTGGAGTGAC TCATACCATG GGATTAAATA CCAGGGATAC AGGCTGTAGG TTCGGTAATG
117241 GGTACCAGGC TGGAGTGACT CATACCATGG GATTAAATAC CAGGGATACA GGCTGTAGGT
117301 TCGGTAATGG GTACCAGGCT GGAGTGACTC ATACCATGGG ATTAAATACC AGGGATACAG
117361 GCTGTAGGTT CGGTAATTTA AACCTATATA AATTTTACCC TGTTGGGGGA ATCCCCGTTT
117421 GTACCTGAAT TAATATAGGA ATACAGGCTG TAGAATCGGT AATGGAAACC AGGCTGTAGG
117481 ACTAGGCTGG AGTGACTCAT ACCATGGAAT TCAATTAACA AGGAAAATAT AATAGAATAT
117541 ATATATATAT AGGGAACTTG TAAACAAAAC CCAACTCGCG GATTGGCTGC CTGTTTGGGC
117601 CAACCAGCAG CGAGAATTCC CTGATTGACA GGCGGACTGG CCAATGGTTT GCGAGCATTT
117661 TTGATTGACA GAACGGCCGG CCAATCGCAA CCGAGAACTC GGCAGCGAAG CAAAAGACAG
117721 ACGGCCGCGG CGACCAATGG CCGCCGCGGG TTAGTTTGAT TGACGGCTTG GCCGGCCAAT
117781 GGGAAACGTC CCTGCCGGCG GCCCCTAATC CCCCTGGCTT AAGGGATTAC GGCCGTAAAC
117841 ATCGCTGGCG CGGTGCCACC GCCGCCGAAC CCCCGCGCCC GGTTCGGCGT GACCTCGCCG
117901 CGACCCCGCC GCCCTCATGC GCACCATGGG TGCGATGGGC ACCGGCGGCA ACATTATTCG
117961 CTGACCGCCG GACGACCCCG GCGCCAAAGA GCGGCCGATG GGTACGCGCG GCGACCATCG
118021 AAACATTTTT CAATAACCCC AGTCTGACCC CGCGGCCCAA ATCTGGGGCC CATGGTGGCG
118081 CGCGGCCCCC ACCGTCCCCA TTTCCCACGG TTTGCTCATG GGCACCGACG GTGCCCGTGG
118141 TGATATGTTT CAGTGACCCT CGGATGACCC GGGCGACGTT TTGCGCACCA TGGTCCCGAA
118201 TTCCCCCATG AGTACATGTT TCGGTGACCC GCCGGCGACC CCGGGCGGCA AAAGCCGCCC
118261 CATGGGCACC GGACGGCACC ATGTGCACAT CTTTCAGTGA CCCGCGGGCG ACCCCGGCGG

118321 CGGCACCGCG GGCGCTGGGG GTGATATGTT CTGATGACCG CCGCCTTACG TTTGCAGCGG
118381 TGTCAGCGGA ACATATTCTC CCCGGCTCCC GAGCCCGGGG CCCGCGGACC CGAGTTGGGT
118441 TATTAAAGTT ACGTTAGGGA TTAATTTTAA TCCACTACCC CCCTTAATGT GCGGCTCGGC
118501 CTGCGGCGAC GCTGACGCCG GCGGGTTCGA CCCATTGACC GCATCCTCGG TGGCCGGGCA
118561 GAGGCCGGCC GGCCAAAGGT GCGATCGCGG GGTCTGACCC ACCCGTGGGA GAAACCCCTC
118621 CGTTCCGTTG TCTTCTGCCT GCGGTCCCTG GCCCCCTGGC CCACCGGGGA TGTCTTGTGA
118681 GCACTTTCCC CGGTGAGAAT AGACAGGAGA GTGCCCGCAA GACATCGACG GGGTTGCCGG
118741 TCGGTGAACG GGAGGCGCGC CGCTTGCGTG CGTGAGGTGT CGCCGGTGGG CCGTATGACC
118801 CCCGCAACGC GAGGGGCCGC TGGCGCGGAA ATTTCCCCAG AACCGGCCGC GCCGCTTGGC
118861 GGCGCGGCTT CCTCCCCCAA CGACCCCTAA TTTGGTTTTT AAGTTGTTAT TTAAAGCTCC
118921 GCTGTGTTTG TGGCACGCGC TTAAGCGTTG CCACCTGTTG CGGAGATCCC CCGTGCGCAT
118981 CGCCGCCCTC CTCGCCTCGG TGCACGCTGC CGAGCGTGCC ACCTGTTTGG TTGTGTCATC
119041 CCATGTGCGC ATCCCCCCAT CTTATTTCCC CGACCAGGCA ACATAAACGT CACGTCGTTT
119101 GTAAGAAATA ATTGCTTTTT TTATTTTCGT TAACCCCGCA ACCGCGATAC AGTCTCTCGC
119161 CTTTCACCCG TTGTTTTGTT ACCAGCCCCT CTTGGCGGCC GGCACCGGCG TTAACCGGGT
119221 GCGTTGGCGG CCACGACCTT GGCCACGTCG CGCCTGGAGA CGGCCTGCAG CAGGTCACTT
119281 AAAACCATGT AGTTTGAGGG GCTGACCAGG GCCGCCTTTT CCATTTGCGA GAGCCACCGA
119341 AAAAAGGTGG GTGACTGGTT GCGCCTGCGA CCAAGCTGGT CGCCCGTTAG AAAAACTAAA
119401 TTTTTCACGT CCCTTTCTGA GAGCTGGCGG TCAATGGAGA GCATCAGGTG TTTGTATGGA
119461 CTAAAGTAAC TGGGGCTGGA CCGGAGCCTC TCGATAATAA ATCCCACGTC TAGTAAAAAA
119521 AAGGTGGAGA CCAGGTCCAA TCGCCCAACG AGAAACACGG CCTCCAGCAA CACCGGGAGG
119581 GGAATACCGG GTTCTCCAGT TGAGGGGCAC AGCCCGTTTG CAAACCCCTC GGGCGTATCG
119641 TCGGACGCCG GCCGGTCAAA AAGCCACAGC ACGGCCTCTC TGTCGTCTGC CTCTAGGTGG
119701 CGACCGAAGT CCACTAACCG CTTATGCGGG AACATGGGTC CGGTGCGGCG CTCGCCGACA
119761 CCCTCTTCTA CCTACACTAG AGCGTTTGAT AAATCACGAC GAACCGCGGG CGTTACTTAA
119821 AATGTGGTCT TCGTCAAATT CCCGCAAGGT GGTGGCAACC GACTCGGCGG CGGCCGCGAG
119881 GGTTGCGGCC GAAACGCCGA CGGCCTCCGC CAATCTGGCC GGAGGCGGCG CAGCGGCACC
119941 GAGGCTAAAC AGCGCGCAGG CGGCCGCCAC CAGGGACGGG GGCAGACCGC CGGTGGCCGG
120001 GTTCACGATG GCTTTGTGAA CAGACTCCAC CACCTGGCTG TGAAGGGCAA AGAGCTGCTC
120061 TTTTGTAAAG CCGCTTTTAA ACAGGGTGGG CCCGATGGCG TCCGTTGGTA AGACGGCCTC
120121 TAGTTTCCAC CTTAAAGCCT TAAGGATTGA TTTTTCCTGT TTCAGTAAAT CCGCCACGGA
120181 AAAATCCTCC GCCGCGCAAA AGCATAAAAA GGCGGCCTTG AAAGGATTCA GGTCTCTAAT
120241 CTTGCCGGCT AAAAATAGAC AGGCGGCCCC CAGCCTCTGA AACCGCCGAC GGGGGATGCT
120301 GCGACACTCT AGATAGCGAT CCAGAATACT AACGGCCAGT GGGAAAACGC TAGCATCCGC
120361 CTGGTGGGCC CTGGCAACAG ACCGCATCCA CGTCCCCAAG ATAACTCGCA TGCCCACCGT
120421 CACCTCCGTT TGTATCGTCC CATAGCAGGC CGTGGATGTT ACAAAGGACG CTTCGTGGGC
120481 CAGGAGGTTA GAGAGGGCCC GGTCCTGGTA CAGGACGGGG TCAATGGTCC CGGTGGGCAC
120541 TGGGCCAACA GAAGCCATTG TCAAGGCAAG GTAAAGAAAG CGATATAAGT CCTACCTGCG
120601 GCTCACGGAG GCTTAATACG AACGTCCGGC TTAACACACA ATCTTCTTTG CGCTCTATGT
120661 CCGCTCCACA CTGGTCCCAT GTCTTCTGCA GTTCAGCCCT GCCTGTTCCC GGAAAGCGGC
120721 GGACTAGATG CAACGCGGAA GCGGCTGTTC ATTTAAGGTG AGTCACAGGC GGGGCAATTT
120781 GCCAACCAGT GACGCGTTTC CTGCAGTCTG GGTTTTCACC TGGCAAGCGC CACTGGTTCT
120841 TGGCAGCCGG CCGGTGGGGA TTTTTTTAGT GCTGAATTGG CAGTCCTCTG TCCATGCGCA
120901 CTATGCTGGC CTGGAAGTGG GGGGCGGATG GGTGATGTCG TTCCCACCGG CGACGGGCCT
120961 TTTTGACGGC CTTGGCGGTT CCGCGGCTTT GACAGAACAC CGTAATGCAA AAAGCCGGGA
121021 TGACGGGACT GCCAGGTGCG TCCGGCAAAA CGCGGACGCT GGATGCTCTA ATGCCGGGAA
121081 ACTGTAGGCA CCTTGCCAGG CGATACAGTG AATAGGGGTC TGAGTTACAA AAGATGACTC
121141 CGTGACACCA CGGGCCCGGC CGAGGGGGGT CGATCCTGGG TCCCAATACC TGACGATAAA
121201 ATCTTCTCTG TGGGCGGTAT TTTCTGGGTA CCTCTTCCTC TGATGAGGGA TACGGTATGG
121261 GAATGGGGTA TCGCGGCGGG CGCGGGGGGC GCGGGGGGCG CGCCGGTTGC TGCTGCCGCT
121321 CCTCCCCCTC TTGCTCCTGT TGTCTTTCAG GAGCTTCGTC CTCGTGGGAG TTGTGACTCG
121381 TGCATCTGAG ACGTAAACAA GGAACTCCTC CGGCACGCGT GGGTGGACCC CAACCCCCTA
121441 CGGTGTACGA GTGGGATGTA TGGCCAGGAA ACGGGGGTCC TGGCCGAGAC GGCCCGGTGG
121501 GGTGGTCTGT GTGGCTCACC GCCTGTTGCG TGTTAGGGGA TGGGGCCTGT TGAGGCGTTG
121561 GCGATTGGGG TGGTCCCGGT GAGTCGGGCG GGGATTGCGG AGGGGGCGAG GACGTCGGCG
121621 TTGGCGATTG TGGCGGTCCA GGTGAGTCGG GAGGGGGCAA GGACGTCGGA GGCTCAGGCG
121681 GTTCAGGTGA GTCGTGGTGC GATGGTTGCA AAAACGACGA AGGCTGCTGA TTGGGCGATG

121741 GATTATGCGA CGACGGCGGT GCAGGTGAGT GGGGGCCGGG GTGTGACGGG GGTGATGGTG
121801 GGGGCGAATG CGATGCGGGG GATGCCGGGG ACGACCCTGA CGTGTTTTGA GGTGACAGGC
121861 CTGAAGCTAA CACCGGCGCA GACCCGCTTG GCGACGGCGT TGGAGATCCC AACACGATGA
121921 CGGACTCAGA CCCTGACCCT GATGGCGACC CTTGGGATGA CCCGTGCCTG GATTCGGTGG
121981 CCGTTTCGGT GTCCGATCCG TGCCCTTGTT CTGCAACCGC CCGCACTTGC GGCCCTGATG
122041 TTGTTGGAGC GGGGGGCGCG GGTGCGGCGG CTGGCGCGGC CGTGCCACGC CCCCTGCCGC
122101 GCCTCCCGGC CTGTTCTCTA GTGCCGCCGC GTCTTCCGCA ATGTCCCCTG CAGGAACTTC
122161 GCAACCCGTG TCCCGATACT ATGCCGGAAC GATGTTGCCG GCTGCCCCAC ATGTTCGCGG
122221 TTTACTGCGT GTAAAAAACG GCAGATTAAG TAGATTGCCC TTTCCACGAC CCGCAGCCAC
122281 CTGAGCGCGC GCACGCAGAA CCTATGTTTA GTGCGATTGC GATTATGTCC GCTAGGTGGC
122341 GGGCTATTTT GATCCTGCCG CGCACGCATT GACCGTTAGG TGGCGCAGAG CGCCGTTAAC
122401 CGCAGCGAGC ACCGCACGGC GGTTTTAGTT TCGTTACTAT AACGTTTAGT TAAAATCCGT
122461 GGCCGATATT GGGTGGCAGT GTGTAAGTCA CGCGCGCCGC GGCCCCCTTA CCTTGTCGCC
122521 GGCGAACCGC GGAAAGTCTG TTCCGCCGCG CAGCGCGCGG CCGCTGCCGA AACCAACGCC
122581 GCTTGAATTT CGCGGGCGGC TGGGCTTGGT GTCAGTGATT CACACAGATG CGATGTGATT
122641 GGGCGTGGTG GTCAACGTCA CTTGTTAACC CGTAAATCTG GAAATTCTTG TCCGCCGCGC
122701 CGGTGAGCCG CGGCCCTAAG CCTTATAGTG CTGTGCAGCG ATCCTGAAAC TCGAAAGCTG
122761 GCCATATAAA CCCCAGTATG GCGTTGCGTG TCGGTGGCAA CCTTTTTGAA AAAGATTTAT
122821 TGCCGCCGGG TGTTAAGCAC CAACACCGAC CGTGCGTTTT CAATTATGTC GGGAGGAATT
122881 ACATTAACGC TGCTGCTGGC GACGCTTGCC ACGGTTCGGT GCGCTCTTCA AACGCACTAT
122941 GCGGCGGTCC CCGTGCACTC TACCGCGTCC CTGGGGTGCG TGTTAACGAC ACCCCACGAC
123001 GTTCTTATCG TTACCTGGCA AAAACAGGAA TCGCCTAGTC CCGTTAACGT GGCCACATAT
123061 AGTTCCGAAG CGGGCACGGT GGTTCAGCCC CCGTTCGCCG GTAGGGTTGA CATTCCCGAA
123121 CACAAGTTGA CCAGAACGAC CCTGAAGTTT TTTAATGCCA CCCTGGAGGA CGAGGGGTGC
123181 TACCTGTGTA TCTTTAACGC GTTTGGAGTG GGAAAGCTGT CGGGAACCGC CTGCTTGACG
123241 GTTTACGTCC CCCTGTCCAT GTCCGTCACG TTTTACCCCC CGATTAACCC GACGCAGCTC
123301 GTCTGTCGGG CCGAGGCCAG TCCCGCACCG TCGGTCAACT GGACCGGCGT GCCGCCCGAG
123361 CTGTGCAGCG AACCTGAAGT GTTTCCCCGG CCCAACGGAA CAACCCTGGT TGTGGGTCGC
123421 TGCAACGTAA CGTCGGTGGA CCCCGAAGAC CTTGAGAACG CCACGTGCCT GGTCACTCAT
123481 ATAGGCGGTT TGGCCGCGGC GCGGCCCCTG GACCCCGTGT TTTCGGATCC CCTGGAAGGG
123541 ACGAGCCACT ACGTGGTGGG TGTGGTGGCA GCGGCCGCCG TTTTAGGCAT TTTTTTAACG
123601 GGTGTTTTTT TGTATAGGTC TATGTGAGCG CGCGTGTCCC CCGTGTCTAG TGTTTTGTTC
123661 CCCAGTGAGT GTCTCCATGA CAAATACAAA TTTGAGGCTG GCTTTTTAGG GTGGTTTCTT
123721 GTGCGACGCT TCCTGTGTAA CTGCATACAC CGGGGTGTCG CCAGGAAACC GCGTCTCCCC
123781 TTTATGTCCG CTCGCCCTCC CAGAGCGAAA GTGAGAATGG TTCCTGGGGC GTTTTGGCGT
123841 TGAGAGAGTC GGGCGATGTT GCCGTAGCGG CGTCTGCAAA GGCTCACCCG CTTCTGTTTT
123901 TTTTCTTTTT GTCAGACAAC AACATGGACG CCTTGAACAA TAACCTTAAC CTGCTGATGG
123961 ATTTTCTGTC TAACTATTCG AATAGCTACA GTAGTTATGA CGACAATATG TCTTACACCT
124021 TAGACACGGA ATCCACGCTG TGTCGGCTGA CGGTGGTTTT CCCACCTACC GTTTATGCGA
124081 TTATATGCTT TTTTATTTTT TGCATTACGC TGTTTGGGAA CGCGTTGGTG CTATATATTT
124141 TTTTTAAATT TAAAGCGCTC GCCAACTCTG TGGATGTACT GATGGCTGGG TTGTGTTGTA
124201 ACTCCCTGTT TCTGTGCGCG TCGTTTTTGT TCAGCTGGCT GCTGTACGTC GCGCCACAGA
124261 TGCTCACGTC CGCGACGTGC AAGGTGGAAA TCTTTTTCTT TTACCTGTAC ACGTACTTTG
124321 GCGTGTACAT TGTGGTGTGT ATCAGCCTTA TCAGGTGCCT GTTAGTTGTG TTTTCCCGCC
124381 GCCCGTGGGT CAAGCACGGG GCCTCCGGCT TTCTCTGCGT GTGTGTGTCT TTAATCGTGG
124441 CGCTGGCGCT GTCTGCCAAC GCGAGCCTCT ATAGGACGGC CCTGCGTCAC CCAGAGACCA
124501 GCGAGTGGAT ATGCTACGAA GATGCCGGGG AAGATACCGT CAACTGGAAG CTGAGAATCA
124561 GAACCACCAG CGCGATCTGC GGGTTTTTGG TTCCGTTTGG GCTGATGGTG CTCTTTTACG
124621 GACTTACGTG GTGTATGGTT AAAAGCACGA AGCTGGCCAG AAAGGGAGCC GTTAGGGGTG
124681 TAATTGTGAC GGTGGTGGTG CTGTTTTTAA TTTTTTGCCT GCCCTATCAC CTGTGCAACT
124741 TTTTTGACAC CCTGTTGAGG ACCGGTTTTC TGGCCGAAAC GTGCTACCTC AGGGACGTGA
124801 TCAGCGTGGC CATGCACATA TGCTCCCTGC TACAGAGCAT GTATAGCGCG TTCGTGCCAG
124861 TCGTGTATTC TGGTCTTGGG TCTCTGTTTA GGAGAAGGGT TAGGGATACC TGGTCCGTGT
124921 TTAGGTGTTT TTCCACTTCA GGTAGTTTAT GAGACACTCA CGCGACACTT GGTTGGATTG
124981 TTTGTGTACA TTTATTTTCA TTTTGTGTAC ATTTATTTTC ATTAAAGCGA TCTGACCTGC
125041 AGACCTTACC TGACGTTTAC TGTCTGTTTC TTATGCACCA GAGGAACAGG GACTGGAAGG
125101 CCAGCGCCCA CGGGGAAACT GTCATGTCCG CAAGCTCGGG CGGCACGTAC GCCCACTGCC

125161 AGGGGTGGAA CGTCACGGCG GGGTCGACCA GGGAGGCCAG GTGGCGCCCG TCCGGTGACG
125221 TGATGGCGGC CACGGCCGAG TTGGCAGACG GGTTGCGTGG GTAGTGCTGC GCGAACATCC
125281 TCGGGTCGGC GTTGCCCGTG TGGTAGTTCA GGGCGATGCG CTGCTGCTGG TTGAGATGGT
125341 ACTCCATGGC GTCGCGCGGG TATCTCACGC CCAGGTACCG GCCGTTGGCC CACCCTGGGA
125401 GGACGAGGCC CCGAAGAACC CTAAACATGA TGCTGATGGT GGTCTGGGGG ATGTGGAGGT
125461 TTAGCCAGAG GCACTCGTGG TTCCCTGATG CGTTCTCCTC CAGGTGGATG TCCCACTGGT
125521 CGGGGGTTTT GGGTCCGGGC GTGTCGTGAG GGGTCTCTCT AAGAAGACCG AGCGCCCCCA
125581 GGAGCTGGAA CCCAAACTCC CCGCAGCATA GTGAAAATGT ATCCGCTCGG CGGAAGAAGG
125641 CCATAAGGCC CCCATAGCAC CCAGGGTCGT TGAGAAGACC CATGATCGCG CATCGGGCCC
125701 CCACGTAGCT GTCTTCGATG CCCACGGTTC CACCGATGGT CAGACCGGAG AATCCCCGGA
125761 GGATGTTCCC TCCTCTAAGG TCGTCGGTGG AGACGGCCGC GACGTCGAAC CCGACGTTGG
125821 TGAAGGCGGC CATCAGCGCC CTGGGAAGCG GGGCACCGGG GGTGACCAAG GCGGCCACTG
125881 CTGGCGGCCT CGACGGCGTT GCAAACAGAG TCAGTTCGCT GTTTCTGCAA ACCTCGGCGA
125941 GGTGGCCCAG GTTGTGTTGG TTGCACCCGT AGTCCTTTCT GTAGAGTTCC TGCGCGGGCG
126001 TGAAGCTGGG CCCCCATGAG TACCACTGTT CGTCCGAGAA CGAGGTCCAG TTTGCCGCCA
126061 CCGAAGTGAG GGTCTGTGAA TAGACTTCAT CGTTGTTGTG TGAGATGACG ATTCTTTCTG
126121 ACAGACCCTC CTGACCCACG GTGCCGCACA CGGTGGCCCG GCAGTCAAAG TTTTGGCACG
126181 CCTGGCGCAC CTCGTCGACG TGCTGGGGCT GGATTTCGAA GATGACGCCC GGGGTTTCGG
126241 ACACCAGCCA CTGCAGGGGC GTTTCCTCTG ACGGGATGCG AATCTGTAGG CCTCGGTTAC
126301 CGGCCAGGGC CATCTCGATA GCGGTAACCA CCATGCCTCC GTCGCTGACG TCGTGGCCGG
126361 ACACCACCAG ACCCCGGGAG AGCAGGGCCT CTGTTAGCAT GAAGAGGTTG GCCAGGTGTG
126421 TCGCCTGAAC GTCCGGTAGG GTGGGACTCG GCAAGAGACA CAGGTGCTCA AAGGTCGATC
126481 CCTGGGTCAG GTGCGGCACG GGGAAGGACA CCAGCACGAT GAGGTTACCG GTGGCCTTCA
126541 GGTCAGGTGT GACGCGTTGC CTGGACGACT TCACCTCGGC GGTGGCCGTG ATGACGACTG
126601 CGTTAAACGG CACGGGGGCT ACCGTCTGAT GCTGTTGGGT GGCGCTGATC AGTTCTTCCG
126661 ACAGGCACGG GGTGCTGCCG GCCGACGTGA TGGCGAAGTT GATGCTCAGG TCCCTGCATA
126721 GTTCCTTGCA GGCAAACAGT GTGTGCTGCA GGAGCCAGGC CTGGTCGTCC TCCGGGTTCC
126781 AACCGACCGA CGCCGTCAGC GTGATGTCGC TTAGACGCCG CACGTGCGCT AGCATGATGT
126841 TGGTGACCGC CTCGCAGATG GCGTATCTCG CGCCCACTGC CGCGTTGATG CCCATCTTGT
126901 ATGCCTGTTC TCCATAAGCC ACGCACTGTC CGACCCACCG GTTCCTGCCC CTGCCCTGGG
126961 TGACCCAGGT TTCTGGGCTA ACCAGGCGTT CCGTCGCGTC CTGCCTGGTG ATCGGCTGGG
127021 GGCGATATAT GGGCACCCGC GCCATTCTGT CCGTAAACAC CGACGTGTGG TTGACAATGT
127081 GGTAATCTGA CAGCGGCCGG CCCAGGGGTC CCACTTCGCA CTGCTGAGCC ACGAGTCCGT
127141 TGGAGCATCT GTCCACGTGT CGCGTGACAA ACTCTTTGCT GCCAACCGTC GGACATCGCA
127201 GAAGCTGATC GACCACGGAG GCCAGGCGGA ATAGGCCCCA GTCGGTCGGC GAGGGTGCCT
127261 GTTCGCCACC ATCCTCAGGT TCTTCTGGGC GCACCAGCCA GTTGGAGGAA ACGGGCTGCT
127321 CGTCGAACCC AAAAGTTGCG ATGACTTCGG CCGATGTTGG CTGTCTGATC GGCCACGCTC
127381 CCGGGTTCCA CAGCTCCAGG TCGTTCACGA ACTGAATTCC AAGCTCTTCG CAGGTTTCCC
127441 CAACGACGGC GAAAGGGCAT CCGCAGATGC GTGCGGCCCT CCTGATGGCA TCCAACGGGC
127501 CCGGTTCCGG CTCCGTGTCG CGCGGGAGGA CGGCAAAGAC CACGCAGGCA GACACGTTGA
127561 GAAACTTGTT AACTAGGATG TCGTTCACGG TTTCCCGTGA GATGTCTGGC GTGGCGGTCA
127621 GTTGGCTCAG AATGGTTGGT GGAAGCTGAG AGACAAAGAG GCGCAGGCCG CCCCGCGAGA
127681 CCAGGGCTCC CAGGTGGCCG ATGACAGAGG CGGGTTCATG GGCGCGATTA ATGCTGGTAA
127741 ATAGGGGCCC TTCCAAAATC TCGTAAGCCA GCGACAACGT TTGGGTGATG CGGGTAATAC
127801 CTGCGGCGCT CCCCAGGTAA AGCGAGCGCT GGTTGCCATC GCCGTCAACG GGTTCAAAAG
127861 ACCCAAGACA GATGATAGGC TGACTGAATA AAGACCTCCT GAGGAATGTG TATGCTTGAG
127921 GCACGCACGA TAACAGCGAT GTGGTGGTAA TAACGTTCGC CTCGCTGGTG GTGGCGCTGT
127981 GGACGGTTTT AAGAAAGCCA CCGAGGACAG GCGTGCCGGT GCGCGCCAGC TGCTGGCAGT
128041 GGAGTGTCGC TGTTGTGTTA ATAGCTCTGT GGCGTGCGCG AGAAAATGCG CCCACTGACG
128101 GCGTAGTGGC GTAAAATCCA CACAGGGCCG ATCTGATTTC GCCCCCGGGC AGCAGCGTGG
128161 CATGGAACAG AATGTTTTGC AGCGCACCTG CACCTGTCTG AATGGCGCTG TGCAGGTGTG
128221 TAAATGAGTC AGCTATGTGG GTTACGTTTT CGCACGGGAG GCGCTGGAAA ACGCAGCTGT
128281 AGCTCCTGCG GGCCATGACG GCGTATCTGG ACACCGTATA GCTGTTGCCG TTGGAGGCCT
128341 GCACGGCCAG CGGTAATATG TTCCGTTGTA ACGGGAGCAT CACCGCGGCG CATATTGGGT
128401 CTCTCTCGGG CAACGGTCTC CAGGCGGTGG ATTGCAATGT CACCGGGAGC TGGCGCTGCA
128461 CGCGCACATC ATATTCTACT AGTGTGCCAA GTGTCAGGGC GGCCAAGGTG CCCGGGGGTA
128521 CCCGATCGGC GACAAACCGT CTCGTCTGGC AGCTCAGATG TCTGCCCACC TCTAACCGCA

128581 ACGTTCTATC GAGGCCGCTG TCTCTAAGGC ACTGCACTAT TTCGCTAGAT CTTGTGGTAG
128641 GTTGGCGCAT CAGGTCTGGG CCATAGGGCA GGGATATAGT ATGATGGCCC TGGAGGTGCG
128701 CCGACCGAGT TTCCAGAGGT GCATAATTAA CCAGAGGCGC CAGAAGATGC CAGATAACGC
128761 CCAGCGCTGG GTCGTCATCG TCGTTGTCAA AGTTATCAGA GGGGCGTGTT ACAAGCCAAA
128821 GTAAAAATTC ACTAAAGCAC TGTTCTGGCA CAAGCGCCAA AACACTTTCT GCGTCCGAAT
128881 TGTCATGAAT AAAAGCCTCC TCTTCTGGAC TAAGGGCTGC AGCGGCCCAG CGAGGGTTCG
128941 TGCGTTGGGC CATATTTGCT AAGTCTAGCC GGCGCAAAAG CAAGCGGGTG GGACTTTAAT
129001 AGGGCGTATA GGGGTTGTTA GGTATGCGGT GGGGAGCCGT CCATGTTACA TGGACTTCAA
129061 TATTAATTTA TGTTAGTTTT GGGCACCTTG GGATGTTACA CTTATAACTT CCCCCATGTC
129121 ACAGTAATCA CTTGACAGAT CCCCTTCTAA CTCATAACAA CTGTTTCGGG GTGTTGAGGT
129181 TACAGGTAAG GGTGGGAGGG GAATGTTAGG TGGCGAACCA AAGCATAACG TCGCGGGGCG
129241 GGAGTGGGAG CCTGTAACCC CAGAGATGTC ACTTACGCTT GTTAGGCGAT CGACAACTAC
129301 CTCCGAAAGT GTGTCCTGTT GGATGTTTTG TGGATTAGGT AAAACAGAAA GTTGTGGAGA
129361 AAGTAACCTT ATAGATGACC TCCTGCTGGC GCGTGGCTGG GTGACAGCTG GTTGTTGATG
129421 AATGAACACT TCCTCGGGCG TTCCTGGGGT GGTGGGTGTG GTCTCACTAA GGACGGCCAA
129481 CTCAACACTA ACTTCCTCCC GTTCAGACCG TGAGCGGTTT CGTCTGTAGC GTCTTGTGGA
129541 GGGGGTTGCC TGAACCCTTG TTGGTGTAAA CAATCCAGGT AACCGGTAAA TGCACGCGGT
129601 CAGGGCGATT TGTAACAAAC CTAAAAACGC ATAATAAGCT TACAAACGCA TTTTTGGTAA
129661 CAGTTTGTAT AATTTAGCCC CAAAGTTTAT ACAAAGCGAA ACTACTTACC GATTAGTAAA
129721 ATGCAAACCG CAAGAAGGAT AACACTCACT GATTCTGTTC CAATTGGTTC GAATTTAAGA
129781 ATGGAGAGTA AATTTGCCCC GCTGAGAATG CCTGGAAATA AGAGGCGCGT ATTTTACATG
129841 TGGTACAGGG TAGAAAACCT TGTAGGTTTC AGTTGAGTCA CCCTGCCACG TAAACCCGCC
129901 TAGAGTGCTG CACGCGGTGT TGGTTGCAGG AAGTACATCC TAATTTTAAC TCTAAAAGGG
129961 GTTCCTTGTA AAATTATTTA AGTTATGGTG TACTTACATA ATATCCAATG AGTAGTTACA
130021 CTCCAAAAAG ATACTCGATA AAATGCGTGA TATGTGAAGT TTGCAACAAA TGCAGCGGTG
130081 AGTAGGCCCA AGAAAATGAT AAACCCTATA TGAGGAAATT TTATTAGTAC GTTTGTGTAA
130141 CTGCAAAATA AAATTTTAAA TAAATACTTT ATATTCATGT TACCAAGTTT AATTTGTGGT
130201 GGTTGTTCTT GAATTATATG TACTCCTAGT TCTGAGGCAA CTGTATGTAG TGCAAGCACG
130261 CTGCCTTGCA ACATAACCAA TAGGCATGTT CCATGATTCG ATATTCGCCA TGCCATTTTA
130321 AGTGCCACGC AGATTGCAGC AATAGCTATA AAAAAACAAA AGTATGTTAA TAGTCTATAA
130381 AGTATAAAAC ATTAAAACAA AATTAAAAAA TTAAATACCT GTATTTGCCG TAACTATAGA
130441 GTATGATAAT AGTAAGGTGT AAGACTGTGT TGTTGTAGTT TCTGTTAATG CATAAAGTCC
130501 CACAAGTGCT AAAATAAAAA AATAAATGTT AGTTATTTAA TTTTTAAAAG TTTATACTAT
130561 TCAAGTAATT CAAAGTTGTG TCATACCTGC GTAAATTGCA AACCAAGATC CAATTATTCT
130621 TAGTATGTTT GTTTTTCCAC ATAGCCACAC ATAACCATAA AACACAATGG GCATTAAAAT
130681 TACTGGCAAC AGCGCAGATA CTTTGTAGTT TGGAGTCCAC ATAAGAAATG CCGTAAGAAT
130741 AACAGTGCCT ATCCATGTTG GTCCCACAAG AAGACATGAT GTTGTTTTAC TAGAAACAAT
130801 TTTCGGTGGT TTTATAACAA ATATTAATAT GCCTGAAAAA TATAAACAAA CGATTAATTT
130861 AAGGCCAATA TCATAGTAAA ACAAACTTTA AACAGGTAAA AAGTTACGTC TTACCTATGT
130921 AAAATTGAGA TACGCATGCA ATAGCACTAC ACGTCTGTAC CAAAATACTT GCGTTGTCCA
130981 AGGCCAATGC CGATAGTAAA ACACAACATG CTGCAAACAA GTACACAAAA TATAAGTTTA
131041 ATTTGCTAAT TAAAATTAAT ATTTCAAATG TATTATAAAA TCATACCTGT AATCCATGGG
131101 AAAAAAATTG AAGGCAACGG AGTTAATTGA CCTGCTCTGC TTCGATTACA ATACAAACTA
131161 ATAGTTTTCC ATAAAAGAGA ACAACACGTA ACAAATGAAC AAATAATTGT GACTGCTTGA
131221 TTAGAAGTGT CCCTGGCATG TTGTACCAAG GCTATTATCC AAAACGTTAA TGTAAGAAGG
131281 TTTGTAAGTA ACCAACTACT CCAAATAAAA AAGGTATAGC TACACGCCAT TTTAAAATAA
131341 TCAAAACTGT ATTATCAAGT TAAGGGGGAG TTTGAGCTTT AAAAGGTTGG TTACATCGTT
131401 TCCAGGCAAA CCTCACACTG AGTCACACAA AGGTTTCTGT GTGCCCCTGT ACGGGAATCT
131461 GCCTTATTGC TGATTTTGTT GGCAACCAAC CAAAGAAATC ATGGAAAAGC AGAGTGGGAA
131521 TATTTGTTTT TGTTCGTTGG CCTGGATGCT TGGAAACATT CCTGTTATTT TGATGTTAGG
131581 CCAAACTGTG GAAAACATTA AGAGTTGTTA TTGAAAATAT TTTTTGGTAA CAAAACCATG
131641 TTTTAGGTCC CAAAAGGAAC ATAAAAATGT TTAATAGAAA CATGCCTAGT ACAGAACATT
131701 TTGTTTGGTA CGTAAACTAA AAATATGTAT GGTAAGGGCT AGGGGTAAGG GCAAGGGGTA
131761 AGGGCTAGGG GTAAGGGCAA GGGGTAAGGG CAAGGGGTAA GGGCAAGGGG TAAGGGCAAG
131821 GGGTAAGGGC AAGGGGTAAG GGCAAGGGGT AAGGGCAAGG GGTAAGGGCA AGGGGTAAGG
131881 GCTAGGGGTA AGGGCTAGGG GTAAGGGCTA GGGGTAAGGG CTAGGGTGAG CTAGGGTGAG
131941 CTAGGGTTAT AGTTATAGTG TAAAATAATA GGCCTACGAT AAAATGTCAA CATATATTTT

```
132001 TTAATTTGTG TTTCATTAAC AGCCATGGTG TATTAGAGTA AAGGTTAACG CTTCAATACA
132061 ATATACAAGT AAAAGTAAAC CACACGAATT TATAACATAT TTACAAAAGC GAACCAAGCT
132121 GGCACATGTA ACTTCACGCT ACCCTAAACG TACAATACTG GGACTAGAAC CCAGAGGTAG
132181 TTAGAATATA CGGTAGTTAC AGAACTTTGC AGTTCCCTTA GGCCAGCAGG GCTCTGCGGT
132241 TAATTAAACA AAGTTTAAAG ATAACTGAAA CTTTAGGAAG TGCGTATGGG TGCAATGTGT
132301 TCCAAATAGG GCAAGGGTTA CATAAACTGT TGCCTAGCGG CCGGGCCCGG AGGCGCCCGG
132361 CCGGCGCCGC CGCCGGGCCG CGGCCGCCAT CTTGCGCCCG GGGCGAGGGT CCCCCGCGCG
132421 CCCCCCGGGC CCGCGCGCCG GGCGCCGCCG GCCTCCCCCG GCTCCCGGCC CTCCGCCCCG
132481 CCGGCCCGGG CCCGCGGCCG CCGCCCCCGG CCGCCCCGCG CCGCGCGCCG GGCCGCTTTC
132541 GGTTCGCGGG GCCGGGGGTC CCGCGGGGGG CCGGGGCCGC CCCCGGGTGC CGCTCCGCCG
132601 GGCCCGGCCG ACTCCCGGGA GCGCCCCGGT CCGGCCGAGC CCGGCGGCGC CGCGCCCCGA
132661 CGCCCCCCGG GCCCGGGGCC CCACAAGCCG CGGCGCGCGG CTCCCGATGC CGGGCGGCCG
132721 CCGCCCGGCA TGGCGGTCCT CCGCCGGCCT CCCCTCCCCC ACGCCGCCCC GAAAGGTGGT
132781 CTCCGCGCCG CCGGGAGGGG GGCCGGGGCC CGGGGCGCCT CGGCGGGGCC CGGCGCGGGG
132841 CGCGACCGAG GGCCCCGGGA GAACGGGGGA TCGGGAAAAC GCGAGGGGAG CGGGGGACAG
132901 GGGACGGCGT GTGCGTGCTT GTGAGACACC GGGTACGGCT GCCTGCCTGC TCGCTGGCCT
132961 GCTTGCTGAG GGGACAGTAG GCCTGCTTGC TCGCTGGCCT GCTTGCTGAG GGGACAGTAG
133021 GCCTGCTTGC TGAGGGGACA GTAGGCCTGC TTGCTCGCTG GCCTGCTTGC TGAGGGGACA
133081 GTAGGGCTGC TGGCTTGCTA GTAGGGCTGC TCGCTGGCCT GCTTGCTTGC TCGCTGGCCT
133141 GCTTGCTTGC TCGCTGGCCT GCTTGCTTGC TCGCTGGCCT GCTTGCTTGC TCGCTGGCCT
133201 GCTTGCTTGC TCGCTGGCCT GCTTGCTTGC TCGCTGGCCT GCTTGCTTGC TCGCTGGCCT
133261 GCTTGCTGAG GGGACAGTAG GGCTGCTTGC TTGCTAAGGG GACGGTACGC CTGCCTGATG
133321 GCTTGATAGT AGGGCTGCTG GGCTGCTAGT AGGGCTGCTG GGCTGCTAGT AGGGCTGCTG
133381 GGCTGCTAGT AGGGCTGCTG GGCTGCTAGT AGGGCTCCTG GGCTGCTAGT AGGGCTGCTG
133441 GGCTGCTAGT AGGGCTCCTG GGCTGCTAGT AGGGCTGCTG GGCTGCTAGT AGGGCTGCTG
133501 GGCTGCTAGT AGGGCTGCTG GGCTGCTAGT AGGGCTGCCT GCTGGCTTGC TTGCTTGCTT
133561 GCTAGTGGGG CCGCTTGCCT GCTACTAGGG CTGCTGTGCA GCTGGGAGAA CAGAGTAGGG
133621 CTGCCGGCCA GCTGCGTGCG AGGGCGTCCG AGGGCCAGAC GAGGACACGG GACCCGGGCC
133681 TCTCCCCCGC CCGGACCGCC GGGCACCCGG CCCAGATCT
```

SEQ ID NO 2
CDS nucleotides 1353-2624
SEQ ID NO 3
MFVLVLFMLLQPVSVELLPAKLTSVPTWCPPHPGDTYLLTCRGTSTARDQRSTQWFRNNTLMRGSNFYGRLVSVTPNATI
SDRYACQTKTTTRSNNIDFRVSSSRLTLQERCSSYGYTYANNTRVLRCYSGGNVTLRNVVFHLNGTAVINGTTTNIHTFV
LTEKTGGTYFCSAFIGNEKFYSQTINVFFTSFTFKPTNDIPNESHFNKTGQIQQTASVQHPENYVVFSVPVFSIGVLTGI
AISLIMCWLFTIRCNENSESSTNSYASQTSYIQPSHNQRSNTNECSRHTYRNAHQEESIEELPNQHTSETDSCCQLVLLE
VKNVAYDGPQENTINEVMEQYDDVVVKNIEQTSYEDNVEHMDYSDTINPNFNYYSGLILEEVDEVFYNELENQYHGLILE
NLDHNEYNHLNELNMIEQYDWLE

SEQ ID NO 4
CDS complement(2692..3258)
/note="dihydrofolate reductase; ORF 2; similar to Kaposi's sarcoma-associated herpesvirus ORF 2"
SEQ ID NO 5
MDITVNCIVAVDEQLGIGKNGTMPWPYLRNEMMYFQKMTSTPSVVGEKNVVIMGKRTWFSIPEKKRPLVNRINIILSREL
REPPHGAHFLARTLDDAFNFYRQYKLKEQLNTVWVIGGKSVYESVLNYKCPLKLYITRIMESFDCDVFFPSINFTEYTML
SEIPGKDTNFEENGIKYKFQVYEKNFNK

SEQ ID NO 6 CDS 3676..5613
/note="complement binding protein; ORF 4; similar to Kaposi's sarcoma-associated herpesvirus ORF 4"
SEQ ID NO 7 /translation="MTFKLFPLFLLHAIMYVHCDENCKPPHFTEYRVKSNTEKDLYSV
GETAELICRPGYVTNTKIITTECLQNGTWSTPNFPCDRKRCPTPADLLNGAVHIHGGD
NALKFGSNISYECNEGYDLIGSNVRFCILQDTENVNWDSNEPVCEIQKCIKPPAVEHG

DYLPNQDVYNYGDAITFKCSLSYTLVGSTTLVCTSNKKWSNSFPTCLMLVCESPQIDN
GYIDIGLSRRYNHGQSITVKCSDGYNIVGPETLTCTNTTWVPPLPKCVLVTNNPSTPM
PETPMPETPTPDYQKINLSTAKTATTPNAFVTTVVSPEKDDVTCVKPHFERFMVKAEN
DKEKYSVGASVELICRPGFTKMQSTVSVECLSNGTWTAPNAKCHRKKCPTPQELLNGE
YIVTSGEDAFKYGTNITYKCNEGYQLLGSMVRICMLKDDLKTVDWEPKAPICDIEKCK
PPPQITNGKYHPVKDFYQYLDTVTFSCNRDFSLVGDEMTTCISNTWNKPFPRCEQITC
SAPNIAHGKLLTGSSSVYKYGQSVTIGCETGFTLIGSEISTCKDSSWDPPLPTCVPAV
SMPSDTPKPETKKPNTPTPEAPKPNTPNVGTHTPFKPPPQNPPIAPPMSKWKRHVVLV
LFASVASLLFVLAALYCCFLK"

SEQ ID NO 8 CDS 6045..9443
/note="ssDNA binding protein; ORF 6; similar to Kaposi's
sarcoma-associated herpesvirus ORF 6"
SEQ ID NO 9 /translation="MASKGNAGQPLEDNQGSRAPIGACGYVYAYSKQDFPFAEASILG
NRPSGSGVFSLPILYGLTVEHEFPLTVKAAYKKVDTTTLAVKVTCFHREVIVFHNASL
FRPVFDGTGLNELCEEARALFGYTQFIEPGPPHSIWNPLECPQLPDKDEMFLGVVVTE
GFKERLWRGCLVPAVFQTQQVQIAGRQAFKVPLYDEDLFAPHGHRMPRFYHKDVSAYL
YDSLFTSIAQALRLKDVTAVIHATEKQFMQDHYKIAKIVQAKQFSTTLPKTTDGSSHM
IVDSVVAELALSYGCMFLECPQDACELLNYDSWPIFDGCDSPEARVNALERWSAEQAV
HVAGQLFAANSVLYLTKVQKQAPRGQKGDVNVYNSFFLQHGLGFLNEATIKENGSEAF
KGVPSNALDGSSFTPYHLAYAASFSPHLLAKLCYYMQFLQHHKSSTNQAFNMVHYVGT
AANSEMCTLCHGNTPATCLNTLFYRLKDRFPAVTTPQRRDPYVVTGTAGTFNDLEILG
NFASFRDREEDGNPADEHPKYTYWQLCQTVTEKLSAIGITEDHDNHVNLITNIQSFLR
VFKGIDSIVDGEVMKFVNSMIKNNFNFREHVKSVHHILQFCCNVYWQAPCAVFLNLYY
KSLLWIIQDICLPYCMIYEQDNPAMGILPSEWLKMHFQTLWTNFKAACLDRGVLTGCE
LKIVHRDMFCDFFDTDAGSNGLMAPFKMQVRIARAMMVVPKSIKIKNRIIFSNTAGSE
AVQSGFVKPTGTRDTYVVAGPYMKFLNSLHRALFPDTKTAALYLWHKISQTNKTPVLK
DVPDDELAELVSYVKTNSLAFEETNVLDVVPDSLMSYARIKLNGAILRACGQIQFYAT
TLHCLTPVLQTIDAEEYPHVLGSAAIATPVAYLAEIRGRTALTVQTTARQPVAATGRL
RPVITVPMVVNKYTGVNGNNNVFHCGNLGYFAGRGVDRNLWPESSPFKKTGVSAMLRK
RHVMMTPIIDRLIKRAAGQTISTFEAESVKRSVQALLEDKDNPNLLKSVILELIRHLG
KGCQDLSSEDVQYYLGDYCMLTDEVLFTLDNIAQSGVPWTIEDAGALIEDRQDADDLQ
FVDSDDIATASCQPPEEQLPTPSAGALLAGKKRKINALLSDLDL"

SEQ ID NO 10 CDS 9468..11528
/note="transport protein; ORF 7; similar to Kaposi's
sarcoma-associated herpesvirus ORF 7"
SEQ ID NO 11 /translation="MARELAALYAQLSALAVDLSLVIFADPRSIDGARILKTKTQIEN
LNRDLLPLLREQNSVETSSLSLEVEHLAKNIEDKLGELERSLRQRYSSREHFETLHLR
PECHYHSTVTFQFYGGGLIDVNMCLINDVELLCKRLGSVFYCIGANEALSGLNRVLTF
LSTLRGISPIPHPDLYVTSVPCVQCLREIELVPNQGSSLLAVLADRHCDHLCKKVRAE
PIHGLFETELSQLGLKVTKRSDATQHGVRSSADQLRESSLAAIQDHNIFKRVSASIME
LSNLIYWNAGQTGLQTGTENECSQMARLLTHEADMHEHRALITPKLSATHFYDCFRPD
PIESLFCGGLFNSIDDTINALSRDCSVTFFQQANYTNVMRKQNELFTRLNSILRQGSA
GSQKPATPSEPRTTTVAATAASDVIKDAQYRKEQYMKKVARDGFKKLTECLQTQSAVL
ANALCMRVWGGVAYGEASELVNHFLLRRRFVALPWEARCRSDQILFENSKYIKNSLYS
QRLSREHVEIITLQFYGLITGPLTRQSDLFPGPANVALAQCFEAAGMLPHHKMLVSEM
IWPQIQPKDWIDQTFNRFYQLPEGDLNAVQKSAWCFIRELVLSVALYNRTWEKTLRIF
SLAREKLSISNLDVKGLTSGLYLTYEQDAPLVLISQNTGWIFKDLYALLYHHLQLSDG
HDDN"

SEQ ID NO 12 CDS 11515..14004
/note="glycoprotein B; ORF 8; similar to Kaposi's
sarcoma-associated herpesvirus ORF 8"
SEQ ID NO 13 /translation="MMITNRTRRLLRAWVVIIAIGTAVGENVTTPKGATTTAKPTPGP
STPTPPENPPRAEAFKFRVCSASATGELFRFNLEKTCPGTEDKTHQEGILMVFKKNIV

PHIFKVRRYRKVATSVTVYRGWTETAVTGKQEVIRPVPQYEINHMDTTYQCFSSMRVN
VNGIVNTYTDRDFTNQTVFLQPVEGLTDNIQRYFSQPVLYTTPGWFPGIYRVRTTVNC
EIVDMIARSAEPYSYFVTALGDTVEVSPFCHNDSTCSVAEKTENGLGARVLTNYTIVD
FATRQPTTETRVFADSGEYTVSWKAEDPKSAVCALTLWKTFPRAIQTTHEASYHFVAN
DVTATFTSPLSQVTNFTGTYPCLNDVIQKTLNATIKKLSDTHATNGSEQYYETEGGLF
LLWQPLTPLSLADEMRELNGTTPAPPTTTSTANRVRRSVGTNEQATDDLAAPQLQFAY
DKLRASINKVLEELSRAWCREQVRDTYMWYELSKINPTSVMTAIYGRPVSAKFVGDAI
SVTDCVAVDQASVSIHKSLRTSTPGICYSRPPVTFRFLNSTTLFKGQLGPRNEIILTD
NQVEACKETCEHYFIASNVTYYYKDYVFVKKINTSEISTLGTFIALNLSFIENIDFRV
IELYSRAEKKLSGSVFDIETMFREYNYYTQRLAGLREDLDNTIDLNRDRLARDLSEIV
ADLGDVGRTVVNVASSVITLFGSIVSGFINFIKSPFGGMLMILVIVAVVLIVFALNRR
TNAIAQAPIRMIYPDIDKMQPSGGKVDQEQIKNILAGMHQLQQEERRRLDEQQRSAPS
LFRRASDGLKRRFRGYKPLENEEAQEYEMSK"

SEQ ID NO 14 CDS 14122..17166
/note="DNA polymerase; ORF 9; similar to Kaposi's
sarcoma-associated herpesvirus ORF 9"

SEQ ID NO 15 /translation="MDFFNPYLGPRGPRPHSHRGTDAPAPAGAGAVQPPPDVCRLIPA
CLRTPGAGGMIPVTIPFPPTYFENGARGDVLLANERSMWTARDRKPVAPDPQDQSITF
HAYDVVETTYAADRCAEVPSRFQTDIIPSGTVLKLLGRTEDGTSVCVNVFRQQVYFYA
KVPAGINVTHILQQALKNTAGRAACGFSTRRVNKRILKTYDVAEHPVTEITLSSGSML
STLSDRLVACGCEVFESNVDAVRRFVLDHGFTTFGWYSCARATPRLAARDARTALEFD
CSWEDLSVQADRSDWPPYRIVAFDIECTGEAGFPCATRDGDAVIQISCVFYTTREGAP
NPPNILFSVGTCDPIPDTDVLEFPSEYDMLVSFFAMIRDFEVDFLTGYNISNFDLPYL
ITRASQVYNLRLNEYTKIKTGSIFEVHEPRGGGGGFMRSVSKIKIAGIVPIDMYQVCR
EKLSLSDYKLDTVARQCLGGKKEDVSYKDIPPLFRSGPGGRAKVGSYCVMDSVLVMDL
LKMFMIHVEISEIAKLAKIQARRVLTDGQQLRVFSCLLEAAARENFILPVPTPEGQGG
YQGATVINPIPGFYDEPVLVVDFASLYPSIIQAHNLCYSTMIHGRDLHLHPNLTPDDY
ETFVLSGGPVHFVKKHKRESLLGRLLTVWLEKRRAIRRTLAACDDPSLKTILDKQQLA
IKVTCNAVYGFTGVASGLLPCINIAETVTLRGRTMLEMSKSYVEALTTEDLRTRLGRE
VTARHGARFRVVYGDTDSLFIACDGYSAEAVSAFCDDLAARITADLFPPPIKLEAEKT
FKCLLLLTKKRYIGVLLNDKMVMKGVDLIRKTACKFVQERCRAILDLVLHDPEVKAAA
RLLCKRPPHAVYEEGLPAGFIKIVEVLNASYLDLRNSVVPIEQLTFSTELSRPVCDYK
TTNLPHLAVYQKLASRCEELPQVHDRIPYVFVDAPGSLKSDLAEHPDYVRQHQIPVAV
DLYFDKLVHGAANILQCLFGNNADTTVAILYNFLNVPYKLFS"

SEQ ID NO 16 CDS 17261..18511
/note="unknown; ORF 10; similar to Kaposi's
sarcoma-associated herpesvirus ORF 10"

SEQ ID NO 17 /translation="MLVNELSVVLGDWEVTFHRGRFSFVNLTRLQTFKGHGGYARVRL
PFSLDQLLHQHFAFGLVTRLKELPPFSDCVALIAPLDSGGDADAARVAPGFVLDSSRP
LTVWVNASGRHTIRFCLLFLKPIDLERAVTYVFGENGGARSEGTPKPTCATESLPGGP
LRVSGEASQTSPHSFVAYFPTANSVACLSLLRLQVRPFSDDAAHRDARISPKYVTFSN
SGGNVCKASVHTLSPSRCKTAQMEIIYAPGDPNAEIVLGQSGPVLPTHTGGRVLGVYA
DAEKTIQPGSSAEVRVQLIFQQGAAARGDLAFLVTGVAPEPLFVVTPALLLSGCTTHL
RLFNPNGTPTTIKRDTLVAAAAPCPVVRLSSADDAPRDLVASPDTGALSINAFTIPVG
FPGVVSAECHVSLRDNGVHERMNH"

SEQ ID NO 18 CDS 18520..19749
/note="unknown; ORF 11; similar to Kaposi's
sarcoma-associated herpesvirus ORF 11"

SEQ ID NO 19 /translation="MGTPVRFFRGEWQTSSLVDNGTPRYSSLVWAATIHDGYLTLVNR
SELCVTERSPCLPACPSIGRLVGKRFPGFAFASATLGDRGTRTVFYAFGHRDNPLDIV
PAVVERADRELVLRVHAPQTTRVSRYGLKVFVAIVTVVRPPGVFLHFPQDRVPIALTD
ACSQEGSRLTSEEPWIKIQGFPVLSDETAHPFLLTQKTKPFTERKFCRLIMDNDQRSA
VNTVYLGKQHVRVTVTRPPETIVTDGPVTATLSLTGNAPIAFRHNPYFELPWSSTTAI
FTPVVYVGLTVCIPPNCSKFVRYGNTYVSAFNRKLTAIISNHAHNGGFRIQDCEWPPN

REIEILVTNVSQAPVYISTGTQLGQAIFVFAPRFGGPAKLRQLLGHRSRALELPGGVT
VDSQKLCRFETMYLFST"

SEQ ID NO 20 CDS complement(19921..20544)
/note="R2; similar to IL-6"
SEQ ID NO 21 /translation="MFPVWFVLFYLSCWAASPTLAPPPTAAGINVLPQWAGNRASLDR
TRGRLSEVGLNIQRWFVYLCHHSTLCRVREYPRIMSFVHFPILMSNVECQRREFRGAE
CMNAMVRGLRAYESYLTRLRMLLDDAPGDADAAAIGSAVTVVLSALDSLIEELPVNNK
IGGAESNEKTVRALGGQSPRDVVLSAFRILEYLQMFLRDGRRAIAMM"

SEQ ID NO 22 CDS complement(20777..21778)
/note="thymidylate synthase; ORF 70; similar to Kaposi's
sarcoma-associated herpesvirus ORF 70"
SEQ ID NO 23 /translation="MIVLVHLGICYVKKIIPVCVAGIAAARLRVFSAPEGAAAVRCAC
RGDHGELQYLAHLDLIIKHGVQREDRTGVGTRSVFGLQARYNLRDEFPLLTTKRVFWR
GVVEELLWFIRGSTDSTELSRRGVKIWDAHGSRAFLAAQGFGDRREGDLGPVYGFQWR
HFGAEYRGADANYEGQGVDQLRYVVDLINRRPHDRRIVMCAWNPADLARMALPPCHVL
CQFYVARGELSCQLYQRSADMGLGVPFNIASYALLTYLIAHVTGLTPGDFVHTLGDAH
VYNNHVDPLLLQLRRTPRPFPRLKILRKVARLEDFTRADLSLEGYDPHPHIEMEMAV"

SEQ ID NO 24 CDS complement(22245..22592)
/note="R3; similar to Kaposi's sarcoma-associated
herpesvirus K4 viral MIP"
SEQ ID NO 25 /translation="MRGLFVCVFFAVFACVVDYAFPMGSMSGPAPELCCLGYVTHLPP
PGLVVSYSHTSSQCSVDAVILNTRRGKKLCANPGDDAVKKLLQAVDKRPKKGRRTRRS
LIDDSEEGLGSGI"

SEQ ID NO 26 CDS 26846..27409
/note="Bcl-2 homolog; ORF 16; similar to Kaposi's
sarcoma-associated herpesvirus ORF 16"
SEQ ID NO 27 /translation="MAAVQGPPPPPEEENENSLPVDVYAIEGIFLYCGLGQAEYLHHP
VFSPIKEFISAFLKDSARLYERLLRHTDYRSLRGLNAIGQGMLQINTDGRHNWGRALA
VLGLGAYVVDKVKDDERLLTFAIAVLPVYAYEALESQWFRSHGEWEGLRNYCERILRH
RRNARRHMCYGVAAGLLALVALFAIRR"

SEQ ID NO 28 CDS complement(27515..29125)
/note="capsid protein; ORF 17; similar to Kaposi's
sarcoma-associated herpesvirus ORF 17"
SEQ ID NO 29 /translation="MTPVYVGGYVDVVSLPKIEKELYLEPSIVATLLPYTDPLPINIE
HVPEAHVGHTIGLFQVTHGIFCLGKLTSHDFLALASRLAGDSRAAQIQLNPMPRDPLL
EMLHTWLPELSLSSLHPEELQDPNHPPAFQHVSLCALGRRRGSIAVYGPDPTWVVSKF
DSLTREEAGKITVNCLDLCERQVTPPEFAAPLETLMAKAIDAGFIRDRTDLLKTDKGV
ARVARSTYLKASQFPCAQHCGNRDTRTMSALPEDNITIPKSTFLTMVQSSLDNMRNQG
HRTYVSAPPSMPATAAYPSWIPPPELTVPSYAPPVAPPFPFQSAFAPQPSPYAATYYS
PTYGYAPAPSRHQKRKRDVELSDEPVFPGEEVGIHKDVMALSKNILDIQADLRDLKRA
ASQTSGAQDADQRPQPPPVQFSWPQTYASAPYLAYQPQWYSGTDTHLHAPQPYQSAQG
IQQTQPPPPQPASHHAGLATQPATPAPAAQESVMSNAIPSASAPRAGACPPLDPECGQ
SARAPVEASAQPAPVSQIQKMFCEELLK"
SEQ ID NO 30 CDS 28998..29897
/note="unknown; ORF 18; similar to Kaposi's
sarcoma-associated herpesvirus ORF 18"
SEQ ID NO 31 /translation="MFIGRGSVYGSRVATIEGSKYSSFSIFGRLTTSTYPPTYTGVML
GRCLREPKEMSAGLRGLMWRVIRCENLNTFLPGELRFLHLVLCEMYNYGLNVYLLKEA
IANTGTRDDIVLGRKVPVEFWKIVYDGLKEMGVSDATLLSETKRGALWLYFNGRPCLL
KGLGDYVFCQLGLSHSVRVVPENLTDGNYLYNLGSVIPCRLLVALSYCLAFWGHADHE
PWVRLFAGKIFILYLIISGHIMPRKSILEQVGTSGYGGFVEAVCRDVRAVHGIPAWDF
ASAAPALTSQQTDYLFAFNNSVV"

SEQ ID NO 32 CDS complement(29905..31548)
/note="tegument protein; ORF 19; similar to Kaposi's sarcoma-associated herpesvirus ORF 19"
SEQ ID NO 33 /translation="MRTSEKCCMRYPRKPARQITATFWAPHPNNVLFIHKPSLIEERR
NAFVMRNQQLALRVHTLRKNLLRLELDNVLQTHQRETEVVMRDLETIQNMVGDLRSPG
RETANAQTSLNPQPKIAPQTHGDAFVVTIAPGDPGFTVNQDLRLELLPSLYMNQNQWL
PQYGPWYSSLTDNAMQRRVFPRDLRGTTNFQNSTSLKLMSAVISTAASITQDFYADVR
NVSDTQAALCLLNGYYCHRTGTPLPPTRNGLWDNLGTKLATLVSHLKQNTKGLGFEFT
YSNPRQRASLAPLNKETKYNADFFTNHVIYATLAQSGLLPGSKNPGTGQPPGPDLVYI
LATTLFSEDVPPFQAYQWNLRAGLSALGCLVLVYVLLELAQITPRSPHRRLNLASLLG
GRFSKVEDPSGSKQYLKKGQLFDFLTENYISPILSRAPDAPTSFLFPGAYLAALEAKA
ISHLKHTRPFVNLTGSRFNEIFDILNQKLTFRDAGSLIQAQTSLRLTAEEGLAAILSH
PSPPGLAHEIMKSQFGVYDDYDRVYFLVLGYLPVATSVV"

SEQ ID NO 34 CDS complement(31043..32095)
/note="unknown; ORF 20; similar to Kaposi's sarcoma-associated herpesvirus ORF 20"
SEQ ID NO 35 /translation="MAFANQCKHVATLEALPASRKRAGTRAHLAVYRRLIKHRSLDDI
LKFLSIRPTLRATKNVKFRIFFEVSLGRRIADCVLTVNSEHQKTCYVIELKTCLSAAV
FPGNAIKISQRWQGLHQLTDSVAYIGRAAPRGHENWSVRPWLLFKNQKTLKTIHTESS
AFPPTFINTTSAALNGFFSQWEDAHVRKMLYEIPTKTSAANYRNFLGPPSKQRSVYSQ
TISDRRKKKRVCDAKSTAGAKGSHAAKKPAPARTRQRAANAPTGNRSGHARPRNNSKH
GRGSAVPGQGNRQCPNITKPATQNRPADTWRRVRCHNSPRRPGIHGKPGSPSGAPAKP
VHEPKPMAATIRAVVQ"

SEQ ID NO 36 CDS 32094..33767
/note="thymidine kinase; ORF 21; similar to Kaposi's sarcoma-associated herpesvirus ORF 21"
SEQ ID NO 37 /translation="MAEGGSGFGDELVRQMRDRKPRWDESSDDTDDVDTESTDLEYDD
VFPVVDTHGLMSPGSQNYDVPTSPSGTPWELLHPDALYAHPRCPPKRAVVPGGGARPK
VSAFSARLQYVGRQSFGDRETRQLTGAQFSSESEHEYAEIPERTTTRPVESGDKRNFT
SGRRGAISGPSSTKPSHGAGLTRKTKTSLSVSLKNLLRIKDDDAKVDVPRPVTVPVHL
MQPHPMTEYRNAFLIYLEGVMGVGKTTLLNSMTGMVPQENVLSCPEPMKFWTCVYSNC
LKEQRSIVKQGTHGKLITSARVYACQSKFALPFRATAAGIGRNLQPWLVGNGSTKPAN
WIVFDRHLLSATVVFPLVHVKYNRLTPDHLFQILSLFSAHDGDVVVLLTLNSSEAHRR
IQSRGRKEEKGITQNYLRQVAWAYHAVFCTWVMMQYLTPEQMVQLCVQTVSIEDICNM
NSRLTHRFLTLTKLHEQSMIPMVAEMLVSVKEHVTLMEVCLGLFKELRKLQILIVDAG
EHLDDACGLWGNIYGQVMSNEAIKPRAVNWPALESYIQTLTKLEGNGAY"

SEQ ID NO 38 CDS 33754..35868
/note="glycoprotein H; ORF 22; similar to Kaposi's sarcoma-associated herpesvirus ORF 22"
SEQ ID NO 39 /translation="MARISFIFFFTIIRCSVTDKYVYDEKSNVELEFNGTIYQINWRN
VSKELTSIVMEDAWYDSLLLEPLSVTLEKRKSLLRSSIVNVHNNDYTFCKSSSDHVIN
LTVDFNYSSLPGFTGNFNVMTHALTQGVLLTKRELFTNSTNIMDLFYAEKINAEMFKI
TFDYSNVIISGIITENWILVSVTNSSVKSNMQCVALLFGVPSTFPALKGYVSYRDLLV
VKNSNYALGVIAPKSYNTLDLAFLPKNFTEMFVSVIDSPLNAIDYLKGKLLAIEAKGA
CQNPSNENDILSFFFEVTAVNFLFIKNLQKQQLVNVGCVVRHVAALESLMHLLRLCYP
TFKLYELNLETLSHIAESQVFNLPANSMLSLSVNDQEVVFSMFKIVYNTPKVGGKILN
EIVYITNYMYTKYSENYQLTNTFRRNVMNMYEVLTTIKLNVTDSSVFYPYILFTSMCN
NVEISYMINQIAKPDDITIFRVFSPCFLSLRFDLDENKLRSDAPQTSKRTGSELAQGA
SGFWRLLHAFHATRINEFSVINCTRLAWKQVTALMPLTNITYVISSVRPDHARVYEVS
EVFLNSAMFVSAVYPNCSHFTPPGTALHIPILYNFSAPRIGCPLCDSIVLSYDENQGL
QTMMYVSNPTVQANLFSPYSPFFDNDNFHIHYLWLMNNGTVVEIRGLYRRHALSAIAL
VFAFIGTMSALYFLFKLFSILA"

SEQ ID NO 40 CDS complement(35865..37073)
/note="unknown; ORF 23; similar to Kaposi's
sarcoma-associated herpesvirus ORF 23"
SEQ ID NO 41 /translation="MIKISDLKARLVGGAVQLSNGEYVCHVVYSSALAAMVGLPGPAV
PLPLLFKKFSTIYSNMMPLYAPKRPELSMLRIMVSPHPYALNSCLCVGTDEGERGVSL
FRDPVIRSSDFEDTPITVNSKLVIASNSLFLHCRPFSVPATVKTPPVTLTNNKQITIN
ELANTTQEYDPNAPPTLCSALPPDNKKLRSILKQPPATSESNVQSDCLLADIFFAMGS
RQPQIGESPITAFNTVTIMQRANNSIMFLPNLKLKPIQHLFLKHVLLQRLGLENILFH
FKMLYANTCKAAGPYQREYFESMLSRVKQRLEDMVFCLNSIESHDFQKDFRVLSRAPQ
RLLTATDKYFLMFPPQNRELAIQVGAEVIESICDGTPLSEVLANLSPRVTIQKETGNN
LLKFYALLTV"

SEQ ID NO 42 CDS complement(37123..39321)
/note="unknown; ORF 24; similar to Kaposi's
sarcoma-associated herpesvirus ORF 24"
SEQ ID NO 43 /translation="MLLQGPVLLPACPATVAANAPSPANSDFKTQLAIFCCLATNNEI
LENVSLEALDRAMQTETTFYACRALRRLVLGEGLYPFIHRQGGIVGKTGNEYAGPGLI
IDDAIGCTFSHIETHTFLPTVFTYELSDTVLVQSDERILRSLYCSPLMVCGVNYQSMF
RILCRYLQIWEFEECFAAFTRTLPEHLIGTCYQNYFKLLEPFKILTLARCPPPCAKLH
LNYLKFNILGFTSDWISHPELHRVQTVIIHNIESNPVLLKNLSKQNKFQDIKVASELI
IDYQNIVNQSLDVNLQVKINKKDPGKKPYKVVVVTPKSTYYLTFPPEVPIFRVAMCMS
VAEHVCHSCDRLYPNTEFLGPGETPRVLEAMFSRIQYAPKDRDYNFIFNADQNPDRYE
QARHDHQTEPLPDMFDPVKHMSLHNFKISVFNTNMVINTKITCWSLAGTFESIIDIPR
LTNNFVMKKFSVKEPSFTVSVFYSDNLCNGAAINVNISGDMLHFMFAMGNLRCFLPVK
HIFPVSIANWNSTLDLHGLENQYIVRRGRRDVFWTTNFPSVVSSKDGCNVSWFKAATA
TISKIYGRPLLKKLSDELNPILSVPYARIDQVKNTIFTTLETRNKAQIQTLHKRFIEC
LVECCSFLRLDLGALNRAARLGTFDFSKRIISHTKSKHECAILGYKKCNLIPKIYVRS
KKIRLDELGRNANFMSFIATTGHAFSNLKPQVIRHTIRRLGLHWRHKAKI"

SEQ ID NO 44 CDS 39323..43459
/note="major capsid protein; ORF 25; similar to Kaposi's
sarcoma-associated herpesvirus ORF 25"
SEQ ID NO 45 /translation="MEAALEVRPFPYMATEANLLRQMKESAASGLFKSFQLLLGKDAR
EGGVQFEGLLGVYTNVIQFVKFLETSLAVACVNTEFKDLKRMTDGKIQFKVSVPTIAY
GDGRRPTKQKQYIIMKACNKHHIGAEIELSTDDIELLFIDRETPLDYTEYAGAVKTIT
ASLQFGVDALERGLVDTVLNVKLRSAPPMFILKTLSDPVYTERGLKKAVKSDMVSMFK
SYLMDNSFFLDKSDIAVKGKQYVLSVLSDMVGAVCHETVFKGTNTYLSASGEPIAGVM
ETTENVMRKLLNMLGQVDGGMSGPASYANYVVRGENLVTAVTYGRVMRTFDQFMKRIV
DRPNAQPSVDDDRDAVADGQDSLAKTPIAAAVIQIGDKLVALESLQRMYNETQFPFPL
NRRMHYTYFFPIGLHMPRPQYSTSATIKGVEHPAEQSVETWIVNKNNVLLSFNYQNAL
KSICHPRMHNPMPCGQALGQAFPDPGHVHRYGQRSEHPPNMNLYGLVYNYYQGKNVAH
VPDVALKATMTTDELLHPTSHETLRLEVHPMFDFFVHQQPGAQAAYRATHRTMVGNIP
QPLAPNEFQNSRGLQFDRAAAVAHVLDQSTMEIIQDTAFDTSYPLLCYVIECLIHGQE
DKFLINSPLIALTIETYWNNAGKLAFINSFPMLRFICVHLGNGSISKDVYAHYRKVFG
ELVVLQQALSKIAGHEVVGRRPASELINCLQDPNLLPPFAYNDVFTNLLRQSSRHPMV
LIGDEGYETENDRDTYINVRGKMEDLVGDMVNIYETRNNADHDGRHVLDVGPFNENEQ
HMAVLEKLFYYVVLPACTNGHVCGMGVDFDNVALALTYNGPVFADVVNPDDEILDHLE
NGTLREMLEASDIHPTVDMIRTLCTSFLTCPFVTQASRVVTQRDPAQLLTTHDDGRYV
SQTVLVNGFAAFAIADRSRDVAETMFYPVPFTKLYSDPLVAATLHPLVANYVTRLPAQ
RVPVAFNVPPALMAEYEEWHKSPMLAYANTCPMTPTSLSTLASMHMKLSAPGFICHAK
HKIHPGFAMTAVRTDEVLAENLLFSARASTSMFLGQPSVMRREVRADAVTFEVNHELA
SLDMALGYSSTITPAHVAAITSDMGVHCQDMFLMFPGDSYQDRTLNDYVKQKAGCQRF
GGPGQIREPVAYVAGVPHSDNIPGLSHGQLATCEIVLTPVTADVTYFQTPNSPRGRAS
CVISCDAYNNESAERLLFDHSIPDSAYEYRTTVNPWASQQGSLGDVLYNSTSRQVAVP
GMYSPCRQFFHKDAILRNNRGLNTLVTEYAARLTGTPATSATDLQYVVVNGTDVFLEQ
PCQFLQEAFPTLAASHRSLLDEYMSNKLTHAPVHMGHYMIEEVAPMKRLLKIGNKVAY

SEQ ID NO 46 CDS 43491..44408
/note="capsid protein; ORF 26; similar to Kaposi's
sarcoma-associated herpesvirus ORF 26"
SEQ ID NO 47 /translation="MALDKSIVVSVTSRLFADEIANLQSKIGCILPLRDAHRLQNIQA
LGLGNLCSRDSAVDFIQAYHYLDKCTLAVLEEVGPNSLRLTRIDPMDNYQIKNAYQPA
FHWDNYSELVVIPPVFGRKDATVSLESNGFDVVFPAVVPEPLAQTVLQKLLLYNIYYR
VAETTPTDVNLAEVTLYTTNITYMGRNYALDVDPVGSSSAMRMLDDLSIYLCVLSALI
PRGCVRLLTSLVRHNKHELVEIFEGVVPPEVQALDLNNVSVADDITRMGALITYLRSL
SSIFNLGRRFHVYAFSSDTNTASCWCAYN"

SEQ ID NO 48 CDS 44433..45242
/note="unknown; ORF 27; similar to Kaposi's
sarcoma-associated herpesvirus ORF 27"
SEQ ID NO 49 /translation="MSIPKIMTVSRDNEGTVCEVAVDNGRHRAMIYYPKTTNLANERA
DVVKEAFDTETPVDIVKQIVNEGLAISKKNCVRLALYLYFYLQYVCFALLLTWQLNPY
MDPPGLVFAVNPMGPKHVTKLPHPAIVAVGCGADAICKNCSVPDIKTELGMVYHNGSS
DSGQRAHYGLALLKAAWLVMGNVCPEPVVRQGAALLGPWNRTEWSDFKSAMAATTFCG
SRGVLWSPIHEKNLCRPTWNDVINTSVFTNESLCPNIPVVPESVIVLNGDA"

SEQ ID NO 50 CDS 45408..45683
/note="unknown; ORF 28; similar to Kaposi's
sarcoma-associated herpesvirus ORF 28"
SEQ ID NO 51 /translation="MTAHTNGVLTTTGFSTSQPESVQVSPFYRVITKPPVMGLFFCVA
MCVIALVWYVMRRVCCKGRVVADSCRDPRQPAYEMLNVRLRPHGTNP"

SEQ ID NO 52 CDS complement(45733..46779)
/note="unknown; ORF 29b; similar to Kaposi's
sarcoma-associated herpesvirus ORF 29b"
SEQ ID NO 53 /translation="MLQKDAKLIFISSSNSSDKSTSFLLNLKDAHEKMLNVVNYVCPD
HKDDFNLQDTVVACPCYRLHIPAYITIDETVRSTTNLFLEGAFSTELMGDAATSAQSM
HKIVSDSSLSQLDLCRVKSTSQDIQGAMKPCLHVYIDPAYTNNTDASGTGIGAVIAVN
HKVIKCILLGVEHFFLRDLTGTAAYQIASCAAALIRAIVTLHPQITHVNVAVEGNSSQ
DAGVAIATVLNEICSVPLSFLHHVDKNTLIRSPIYMLGPEKAKAFESFIYALNSGTFS
ASQTVVSHTIKLSFDPVAYLIDQIKAIRCIPLKDGGHTYCAKQKTMSDDVLVAAVMAH
YMATNDKFVFKSLE"

SEQ ID NO 54 CDS 46905..47135
/note="unknown; ORF 30; similar to Kaposi's
sarcoma-associated herpesvirus ORF 30"
SEQ ID NO 55 /translation="MENDTPKDKISEADFQQCQAFFHRPIRDLISSGADALNHFSLSE
SDGHKLERIVLLLDLVGTECLSYTTIAAKNVK"

SEQ ID NO 56 CDS 47093..47746
/note="unknown; ORF 31; similar to Kaposi's
sarcoma-associated herpesvirus ORF 31"
SEQ ID NO 57 /translation="MSLLYHDRCKECQMTRVNSPICRFHNVSNLYQCLDCKRYHVCDG
GRNCVIVYTRENLVCDLTGNCVLDNVQDVCSYGPPERRVPDAFIDPLVSHGTRECLKS
DILRYFETVGVKSEAYSTVVKNGQLNGIIGRLIDATFNECLPVMSDGEGGRDLAASIY
IHIIISIYSTKTVYDNLLFKCTRNKKYDHIVKTIRAQWMRMVSTGDPSRVSATGCFT"

SEQ ID NO 58 CDS 47683..49077
/note="unknown; ORF 32; similar to Kaposi's
sarcoma-associated herpesvirus ORF 32"
SEQ ID NO 59 /translation="MDAHGLNRRSVAGQCDGLFHVILPRGFILANNITCGERQRFFAH
TWFAASGRTSKTLYVWGRVFQNTDPGRGDGPSGPWSGLAISLPLFTTNGKFHPFDVVI

LKADTPDSGSSWTVKFLYMSLIAAYRNAMRGLKDKVSQCTDAAVDGEVHPLTVLKEAL
VSPDTATRPVSACNPLQMLTGLLQSRVRDDYVTHHRALERPGNVRGQVIAPTRTEMPN
GSPSRVRLGFRPPKQANYPKTWAQARHVFSSRAYYVCVYDNEELDTKWQRQDPRPLPL
DWSDPVAYLLEGDLFLGAKQNAFVDSIEKTCRCQNYTIKQFFPVLINRDNETVDLIKE
HFIEACFVIRNQVSERSAWVKAALFRNDSNTYWKDVLGLWEHGPHKLGTAIKLPTSEP
CNADVNWSWLLCDEDITRSISGQSTVCLVVSPTLTAWLVLPGGFVIKGRYDLSSEDLM
FVASRYGHPASSHS"

SEQ ID NO 60 CDS 49049..50059
/note="unknown; ORF 33; similar to Kaposi's
sarcoma-associated herpesvirus ORF 33"

SEQ ID NO 61 /translation="MATQRRHILKSFLNKECIWLRHPGTSAFVRVYTATTAHSAVFDP
PVTSENAMSLNFLNVMIVIMKPKEFGPCVTVYMNGDILDFCATESVAIRDVPGRADLC
LIRFGTLSNAPRSVPIPGPLNPHPRETVPGLTKQEIIYTSQTVPRGQIPDAIKGKEFH
QINPFLWFDGGAFWQLFLSVDFMLLCPALDTVPSLARIVGLLTQCDKSTCKICTGAHV
HVNPYRGYTPPDSQGTSPSCPCLISCGARRAADVLVTGHVNLLGLLFDPKASPKVTKL
RLKRNPRPVPIEDAMSGVTAEGTEVQPTSLPWALIRLPDLASRVMLYGCQNLKSICLR
SY"

SEQ ID NO 62 CDS complement(49977..50960)
/note="unknown; ORF 29a; similar to Kaposi's
sarcoma-associated herpesvirus ORF 29a"

SEQ ID NO 63 /translation="MLLTSYRERLQNNLRVVTDGGCENWFRQPPVIISGNDKTERMAH
PCLGVIHAVNAYSSVLDDYLQTYRRVQEPMPAPTLGKPRISSHATLPRLTEELTNYLK
QTCCRVQMANAKDQYMEYQSAQRTHEAFLECPVYAELRQFLANLSSFLNGSYVPGVCC
LEPFQQQLIMHTFYFIASIKAPEKTHQLFATFKQHFGLFETTDDVLQTFKQKASVFVI
PRRHGKTWIVVAIISVLLSSVENVHVGYVAHQKHVANAVFSEVIATLSRWFPAKNLNI
KKENGTIVYASPGRRPSSLMCATCFNKNVSRCFLSSGSRIASRDWLNPAGE"

SEQ ID NO 64 CDS 50959..51942
/note="unknown; ORF 34; similar to Kaposi's
sarcoma-associated herpesvirus ORF 34"

SEQ ID NO 65 /translation="MFPSSFLNNGHPETERRFVKGVQLALDLCDNTPGQFKLVETPLN
SFLLVSNVLPESRPVRDCPQPEGFDFEHIHLPKLTRMQRVLGRYCDHVNNDDTCVNVK
ASSSNSQGALFYLPYGQDEWNWALTLRKDKLVKMAVEGLSNPTTWKGLEPVDPLPLIW
LLFYGSRSFCREPECLYERNFGMKGPILLPPHMYAPQKDVMTFVHHVIKYVKFLYVNA
GGGLETEPSPPFEASRLRAAIARLGDVEADDAYLSAKCMLCHLYKQNDTISIHETHVG
GVIALGGDGARYITSSVRAQRCTSRGDFVLIPLYNIEGLVSMIREHGLGSS"

SEQ ID NO 66 CDS 51923..52372
/note="unknown; ORF 35; similar to Kaposi's
sarcoma-associated herpesvirus ORF 35"

SEQ ID NO 67 /translation="MASAAAKKMLIKSELESEINKKLSISVFDRFGADSAVFNAQYKG
TRESLRSYNSLKKKDDLATVVGTLETSLREKQSELGLLKGFNRKKIEEFDAVADAVRD
LKDELYGELEILGTLDNESVPVEEESPKDDIIRWKLERLPRVCPKSP"

SEQ ID NO 68 CDS 52278..53585
/note="kinase; ORF 36; similar to Kaposi's
sarcoma-associated herpesvirus ORF 36"

SEQ ID NO 69 /translation="MNLFPWKKSPPRTTLLGGNWSVCPECAPKALDPIPKVQTDVDRT
ASSHITVIKTRKTIAQLKIPNNWGQCSHQATDWTAVLGRGSYGVVRSMSLGRCVKHFG
SRREFFYECIFNDIVRACREKHPLNRGGDRILCFLEPCVPCRALIFPQLTGNLLNADL
KHVNPERLAVEFSELREGVSFLNNICGIVHCDISPENILIKGELTTAYGRLMIGDLGS
ASLHTGTPWTGVMVTSKLGFVQHTYHFKAPARFICKHIYRPSCLLYRCLLSCAGGPQA
HMLNQPFQITPQLGLTIDISSLGYSLLACLEKYLQPADPFPQQGALADASSESAHPLF
YLRCMVPRVVIAEIFSVAWDVPLDLGIDSSGHAPAIPLREAYRRFFANQCSLYRAQYK
EDALENASSRLCNSKLKLVLQKLLVRDYFSHCGNCGDHGFFLR"

SEQ ID NO 70 CDS 53566..55008
/note="alkaline exonuclease; ORF 37; similar to Kaposi's sarcoma-associated herpesvirus ORF 37"

SEQ ID NO 71 /translation="MDFFSDEPMVQEMALLDIDEQQRLLSKMSLANFLKHERVRAFFS
DNKKEISMPAIRFVYNFYLFAKVGDFIGNTDVYDFYVTCVFRGRRLTRLSEVYDACLN
MHPHDRHHVCALIEQVTRGQNINPLWDALRDGIISSSKFHWAIKQQNSSKKIFNPWPI
VNNHFVAGPLAFGLRCEEVVKKILATLLHPGEAHCENYGFMQSPLNGVFGVSLDFGIN
VRSDPKDGLEFHPDCKIYEIKCRFKYTFSKMECDPIYAAYAKLYQKPSMQTLKGFLYS
ISKPAIEFVGEDRLPSESDYLVAYDKEWEVCPRKKRRLTAVHHLVKKCMIHNSTAPSD
VYILSDPQETGGQINIKAHLSANLFINVRHPYYYQVLLQSLVVQEYISLSKGTKNLGT
QKNFIATGFFRKRQFQDPSCCTIGEFAPLDPHVEIPTLLIVTPVYFPSVAKHQLVKQA
TEFWAASAREAFPELPWDLSSLCANAPPTP"

SEQ ID NO 72 CDS 54963..55172
/note="unknown; ORF 38; similar to Kaposi's sarcoma-associated herpesvirus ORF 38"

SEQ ID NO 73 /translation="MGFILSVCKRPTNTVDVKGEPIDVSKEFDPIIGEESIVLLTADG
TAPAALYKPKTKPSKHKNNKLSDFV"

SEQ ID NO 74 CDS complement(55255..56391)
/note="glycoprotein M; ORF 39; similar to Kaposi's sarcoma-associated herpesvirus ORF 39"

SEQ ID NO 75 /translation="MKISRSDSFILSSWVKLLVILGLMFIMSAVVPLTATFPGLGFPC
YFNTLVNYSALNLTVRSSAKHLTPTLFLEAPEMFVYISWAFLVDGYLLCYYAWAILAI
FKAKRVHATTMTSLQTWIVLIGSHSVVFMSILRLWTIQLFIHVLSYKHILLASFVYCI
HFCLSFTHVQAMISCNSATWSLRVLEQQIPENSLLDTLLRYGKPIGANLYLSLIAMEM
LVFSLGTMMAIGNSFYMLVSDIVFGSINLFFVLTIAWYINTELFLVKYLKHQIGFYVG
VFVSYLILLLPVVRYDKVFISASLHKVIAVNISMIPITCILAIILRIIRNDWKWCAKS
PEYAPLPQGPKEKTTKVKYSPELNALYETEEDVSDYEDAYPKYI"

SEQ ID NO 76 CDS 56526..57932
/note="helicase/primase; ORF 40; similar to Kaposi's sarcoma-associated herpesvirus ORF 40"

SEQ ID NO 77 /translation="MNAREVALTGHVLHISLHSTHEREKLIIWQVHLLVCQQCGIQGD
AAYLFVTETLSNTDWGNIPAINRHAPSINEHGRNYMQWELRTRLRNPIIQLLSRQPGA
VNVRVSEPNMVIVGCERALDHSCSVRVTGAYLHCDTTMDFSLDSVVSPTREFWFSEMF
SHCLVSNIEVYLKTTGGLYYRASSATQCRKRAKDGALGILDIFNCESREIQVAGQKYT
LSIATATFHVLWVDEACMWNGALAEFFRALHNKLFGDREGVAPTLTYVCPGATPEGTP
FPPYFSAFPHLPLVFGRPRRLDVTAVQELPKAQIAVHWPPFKDSILGDQLLIPGISPK
KPGTVPVRWPLWVEDVNLSLCETTESVARIVDPHSIVIIKFSSLLCQHLKCHRAFVKN
ELEYIATICSSDLRLFIQEEYNRLLATIFTWAAASGYTWAAIDKTTVFIKAPQLSAAV
SGFCPSLNSCRRKQCYEG"

SEQ ID NO 78 CDS 57917..58528
/note="helicase/primase; ORF 41; similar to Kaposi's sarcoma-associated herpesvirus ORF 41"

SEQ ID NO 79 /translation="MLRRLKITVHFLSQEQQKVVTRLEAHLGLPVQETSHPPDWLKCE
VCSASVFLKIPAGVLYAGLARDPTREAKRDSWLDCLVEGATLLLNNSVLPIGALAGIL
PTLFANRRCVNFWLLPRAWVKSAPICPPLPIDCVTPPQFVVTKRGPICWYKEWPLPVD
VDFMYYLQEALCVFSVVSNGEGTESHADNIRQLEKFEKVLCLF"

SEQ ID NO 80 CDS complement(58525..59343)
/note="unknown; ORF 42; similar to Kaposi's sarcoma-associated herpesvirus ORF 42"

SEQ ID NO 81 /translation="MDQILKRLMGEQHRSEAVMPETECSSRGPYNYPVFPRLMLEVHK
KNSICMASNTPKLCVRGRLNVPDLGVHVRTRLQSATFTGFVFACVVEHEDMIDALDIY

PHVFSDRVQLFKPASASVTELCCILSMLENYDKPPLSFILSALDRARYLHERYTCNDS
AFVLYGIEVIASTLAAYHELNPPQGILRVPPLVRFKLHKLLDENADDMKGLLKPIYLE
SFRLTENVGEEEGHAETFNIFYCGTIFTRHLHNASVLKYFQITSLHSIPRQTLF"

SEQ ID NO 82 CDS complement(59297..61027)
/note="capsid protein; ORF 43; similar to Kaposi's
sarcoma-associated herpesvirus ORF 43"
SEQ ID NO 83 /translation="MFKMNPGFGSTCLVHPTELSISLFEILQGKYAYVRGQTLHSSLR
NPGIFGRQLFIHLYKTALGSCTYDNVLKDWTNFETTLKTRWRGVEHLTPEFKRSTFES
WARTVRLTVDQLLLNTINQVLHTRTVLSYERYVDWVVALGLVPIVRRTPDGDTIARIQ
AHCQQMRKTYASGDVTISRIVDKLAQEITSIMTDVTSIYIPDYAEVSVEFNGDKAAYL
GTYRQKDITVEVVSRPIIYNGRVAFDSPLYRLFTAIMTCHRTAEHAKLCQLLNTAPLK
ALVGSTCNDMYKDILARLEQSSQKTDPKRELLNLLIKLAENKTVSGITDVVEDFVTDV
SQNIVDKNKLFGTGTESTTQGLRKQVSNTVFKCLTNQINEQFDTISNLEKERDDYVKK
IQCIETQLLQSLPEGGRPRHDINILTQNTLQALSGLRDPTINLSECHIPKGSSVVNSF
FSQYVPPFMEMLKELTSLWEGEMFQTYNLTPVVDNQGQRTSIAYSQDTVSILLGPFTY
IIAKLTHMDLINHSLISLSLHDIADQLYVDSRLSVYINDIGHKYCEQISQPGTDGPNT
EASNGGAAPI"

SEQ ID NO 84 CDS 60966..63338
/note="helicase/primase; ORF 44; similar to Kaposi's
sarcoma-associated herpesvirus ORF 44"
SEQ ID NO 85 /translation="MESSVGWTKHVEPNPGFILNMTSDAKVRGVVDHVSRLSNITTSP
PEMGWYDLAFDPAEDSGPFLPFTVYLITGTAGAGKSTSISALYQNLNCLITGATTIAA
QNLSRRLKTFCPTIFSAFGFKSRHINIAVRKAHQTGAVSIEQIQQQELSKYWPVIVDI
MKEVMAKKPNGMYGTISNANFETLSRMTGPCLWTSNIIVIDEAGTLSSYILTTVVFFY
WFLNSWLNTPLYRQGAVPCIVCVGSPTQTNAFQSTYNHGTQKTEISSCENILTFMIGK
KVVSEYVHLERNWALFINNKRCTDLQFGHLLKILEYNLPIPDEVMSYVDRFVVPKSKI
MDPLEYIGWTRLFLSHSEVKAYLTNLHTCLTLGGDTRDTKLFTCPVVCEVFVKPFEEY
KRAVNLTNLTVTEWVTKNLFKLSNYSQFVDQDMSIVATESTERSTQVTFITKFVKNSH
VSLNGKTKKCICGFQGTYFEFKRILDSELFVETHSQDRPEYVYGFLNTLLYNAMYSFH
AYGVTRSHEKYLQDLKFAPLPAALATGRVDLQTVREELNLEDDIFYHVCSPPPPAGIT
SLQVLVDTYCALKDVFASRIKVACRWFGGEFEKETFSAFTVNMVVRDGVDFVSPSERL
NGLLAFASTVESYKIKGYTFLPVAFGRCQGLPLSDDLRKKMPSLVVQDSSGFIACLEN
NITKLTETMEDGSVFQVCCAGDYGVSSNLAMTIVKAQGMSLERVAVVFGSHKNVQTSH
VYVAISRAVNSNYLVMDSNPLKTLLREPVDNTSAKHIVRALHNPNTTLIY"

SEQ ID NO 86 CDS complement(63379..64437)
/note="unknown; ORF 45; similar to Kaposi's
sarcoma-associated herpesvirus ORF 45"
SEQ ID NO 87 /translation="MAMFLSDPSRTPPATPRMLPIPGAPRKKRTRRFLFAGSRTGLPV
PPGYGGPPVIDMTAPEDVFDQDSPPTTPKTPDETDSHSENSDYSDVDEEDEPPVSSPP
RIDPHARDGESFNQSGRLPTVITSTGATTPPSAPAPLTAFGGPRPVAVVTGQHRAPQS
SESDSEDDFFIDDYEDTDESGGEADGFSPRASPAWSGDISRSPAEGGWSSDEEEPVVA
GSTAGQETIIISDDDEVDDRGSVETWDESDADEGTGATDVIDLCSSSDSDDDADHVTS
GGVRAACKRRASRRDCNGDDDVIYVGTTQGPKRRMTSTTGGGATSNPEGPGVSGRQTM
AATPPVCGNDNYPWPWLD"

SEQ ID NO 88 CDS complement(64479..65246)
/note="uracil DNA glucosidase; ORF 46; similar to Kaposi's
sarcoma-associated herpesvirus ORF 46"
SEQ ID NO 89 /translation="MEGWLKTIVWSKMSPEVLEEPSTQTLLLSDSWLEFLNLSPFLKQ
KLAALLKRVMDMSNVTVIYPPIDRIMWWSYCCEPEDIKVVILGQDPYHRGQATGLAFS
VAPDYSIPPSLKNIFKEIANTVPGFTAPSHGCLDCWAKRGVLLLNTILTVERGKAGSH
ANLGWDWFTSYIISCLSAKLQRCVFMLWGRKAIDKATLINGQRHLVLKARHPSPLATA
HAATGSPWPQFLGCNHFKLANDYLVQNRRGAVDWNIN"

SEQ ID NO 90 CDS complement(65222..65731)
/note="glycoprotein L; ORF 47; similar to Kaposi's sarcoma-associated herpesvirus ORF 47"

SEQ ID NO 91 /translation="MRSMYTLSLFITCGFFLITCCTGLVVNPCCKIIPLSDFIFPEPF
EIASFHLTNLALCPGLCTATLRYKADRSTTEICVNGFHLRAFFIRILYKLNYSVPREE
LQLLNYMQYSLDEFLAEFEDFHINGSESGTAYTRPPLLDFSDRSTKVSRIRKVITRRG
DLWRVGLKQ"

SEQ ID NO 92 CDS complement(65999..67168)
/note="unknown; ORF 48; similar to Kaposi's sarcoma-associated herpesvirus ORF 48"

SEQ ID NO 93 /translation="MAVSIPVKGVNRETESNWRSIVTTFERHGNADRAIRSLLRFFKG
VDHPGFLASLVILKDVTIDSEKTIERTDLIPLLQGVRFVTQQIYMHLKDHASESPMAE
IWRDCKERFCLALELACGCQRCASAARQLRACQQACRPPKLNPHKQQCVAARLLTAVY
NQMVLRTRVSVSEFCLNALMCVPREFGFVSGDVRVETSRVASCLNLSWLYLILDSYVR
TDLTNLEMAMSRACRIHGLSAGDPFYSALVWLKNSYACDTNTFFFTVNSTSVTTPILM
DICASLTGPVPDVIKINMLPLVNDQMHPSVCVERANFTGSCPKVSPTHHLDGLKLETT
SLTLAADSLDDILQALELICGDDEGILDSYISDINTETEVDESSIEEEIVFEELS"

SEQ ID NO 94 CDS complement(67398..68303)
/note="unknown; ORF 49; similar to Kaposi's sarcoma-associated herpesvirus ORF 49"

SEQ ID NO 95 /translation="MSRHYGKDHLLNHMYKFHYPPLGMIVGEMNTLTVNARNPLYQAA
TLRVERALYLSKILQVLMQHRQGERFIVPQCRSNMVYCLKELHKITNDRIRGLINSVL
PLVDAGCVGFDEELVRVLPEILKLEYPHAHELLPPHDPTSPLSWCLSHMVGVTKTFKG
EVKEMIDTFHDLSVPSFQYLASLVKKFFLVEEVIYEDYQDTQFNVFLNLCFFWTTVIK
MYQSCIFKDKLLDTIKACIELLKGEARQFFGWYDLNTPNLGSSALVKYTEHLIRALSV
DSSAIPIGEICSHLHHCKHALLNLE"

SEQ ID NO 96 CDS 68494..70038
/note="transactivator; ORF 50; similar to Kaposi's sarcoma-associated herpesvirus ORF 50"

SEQ ID NO 97 /translation="MECASLGPISGLIADLNLFNLFCLYRGSRVKTRGAATCNVPCAE
CAQGIVRILTERALCCTEKMFIASACSGVVIPPQLARVLHDVYAEMKAKCLGAWRRLI
CCRRPIMAIADSVLVTYNTLDAEGKLELRLKALCKLVFQPIFLQRILAPMQLLANGKM
VPDNYFTITGTAEKRRPVVTGSTSGMTCPGSSLVPDSLILPVCEPGLLPAPLVDLSNV
LENPEIILSAPPLSQFVITNTHPSLPQSVSIITPTQGVVPGQCFMDTWKAVSQSIHHQ
AQTPILAAALTGSTSAAPGPHIACSPVAGTSRQVEGSAGVDCGKPACVPQPALPPNVP
AKRMETVAQLGNAPVKNVHIGGRVYAPLVNIPIIDLTSPSGSGQSPADIANTPESRMA
AGSPPFAETAATVPAKRKQPREDVADKRLKGDVRGAATVNHPFPGPSGMRVREQGLFD
LIESSTDVTANASGPKNDDDMLAAILQDLYGLQSPPAIDSPSSNSDNEEIFPEVSPPS
SGHGSP"

SEQ ID NO 98 CDS 70355..70888
/note="R4"

SEQ ID NO 99 /translation="MPRVKTQPKRPQVLEFMPLDLHGGTHTEMDSQNLCPDGQDLLGS
YIYTENNGPFSQIMHNGQSNTGTGESFGSYAAGDGFLGGSVSGMYGNNTGEGACSKRP
SACRKRSAALIHAASEASVAEQGTSQGAHAVSDRIGRDGGADNRLLKVSARLSDKTKS
ALRSHPCLRCYSLMFNT"

SEQ ID NO 100 CDS 71468..72160
/note="R5"

SEQ ID NO 101 /translation="MGFGNIRLGWRLCFMVWVAXIARGRSVCPTWHLTDGKYEAVYRH
YLEECRKHEGSGSLDGSGQTKGSGTKATTEANISIRPNVVTSGQNKEPPGTAPRAESS
HDLPRIKQVNALRLSTPELAQPLPVVKSTPRESQSGGTPWNARPHAFIMHTNDMLNPS

VVLSFRAIRARSTRDTEQSVRDRNTVTTSYRTPGRPSLFQARPSSHGARLPPSPRTMA
RYAESRTICDQN"

SEQ ID NO 102 CDS complement(72401..72820)
/note="unknown; ORF 52; similar to Kaposi's
sarcoma-associated herpesvirus ORF 52"
SEQ ID NO 103 /translation="MSSTRPKTRAPKKELTMEELAAQVQKLSVENKQLKKLINSGDPT
RSGSDPVISNTEKEAKIAAAVSALCNVATRKIEAKVRAATAKAVTRGQMEDALAGISI
RVDVSMDETTRGGIAASADGALRRRRAQSRTRNNDAD"

SEQ ID NO 104 CDS complement(72884..73198)
/note="unknown; ORF 53; similar to Kaposi's
sarcoma-associated herpesvirus ORF 53"
SEQ ID NO 105 /translation="MTGSIVLALALLACLYLCLPVCATVTTSSTTGTGTPPVTTNPSA
APSVTPSFYDYDCSADTYQPVLSSFSSIWAVINSVLVAVATFLYLTYMCFFKFVETVA
HE"

SEQ ID NO 106 CDS 73274..74146
/note="dUTPase; ORF 54; similar to Kaposi's
sarcoma-associated herpesvirus ORF 54"
SEQ ID NO 107 /translation="MAEVTAHTVPYAFDSCKFEIIPKNNSSRIALRNKFPVVVKPGEP
LVVPLGLKIIRAPQCAFFLSGAPTDEVYYHTGLIDQGYRGEIKLIVLNKTKQVVTLYR
GEVNVSLIAFMYASPGPLKCPILNLPHYSLDAGFDVTSPHAMTIPPTDRTPFTLSLYY
KSPQLSTPHVPLIVGRSGLATKGLTVDATKWTQSLVHLRFYNFTKEPIDIPANSRICQ
VVFIHEDHVPSGWNILRSRVQLGSTLQISWAKIRFTDVATLPKTHPLNSRHTQSQTEP
ETARGAKGLGSSGL"

SEQ ID NO 108 CDS complement(74207..74839)
/note="unknown; ORF 55; similar to Kaposi's
sarcoma-associated herpesvirus ORF 55"
SEQ ID NO 109 /translation="MAAPGSFWTCCGFSPFGRVGCQYRPLPDPLNECPTHWRTEIAMG
LPPGVDMGDVKQAEMCTAALRQTYLLAVQSNKITEYLRRFDAARVPAGCQETVRIQIS
KLKSIQNVIWNAMLSLAIGDITVDESAFHALLNKRADETVSLLEMEKLATTIASDDSV
TWAAEINNVLVDTEASSNPSHPVIRQPTPQLAVADNIVPDKIIQDAQADG"

SEQ ID NO 110 CDS 74851..77337
/note="DNA replication protein; ORF 56; similar to
Kaposi's sarcoma-associated herpesvirus ORF 56"
SEQ ID NO 111 /translation="MVDEIRAIFSTSGDMAEVITDILTETQATASFFCVLHDRGDAPI
NTPHAVIKLCLPAKRPGGGPRCLPLMVLNLPAWQVNLFLTGDAPLTSDNIKDRIDLAQ
TEEILEPILSVLACKRSAQQTKHDSFKSKVAWFRAKFVSALRKVYKMTPSPYWMITLL
GSFEASFVLAGTFYFFQSHICTAETLVHLTRLFSSSQGQSLVTVNTYDELGRVFGRSD
FLGIVPNFWAYLKYKMQQDDVESRAIDQTINSIRGGLMLSPQDLVHFIYLSFYECMNA
QTFLSYSRTTSSLPTPATVNPPQLCRRLEADFKEHVMAYYNKASYLSTYITILTVPAP
LPDGYENFQELACQYWCGQSRDVAEIMTRINDQYPQLNLTKDLSGLLDLAALDQYSGG
PKENLFTVASRIPTYRCEFLNKQYFVLMHADCIDAYWKQNIIVPEDAQLQGLTDQDLT
SRIFYCDLGLSLPTFKQQILVSRHEYFNPRLPVYRWVLDFDLKVTEGRRTLNDIYNIC
VTLRQVILETLQLIGPLKPNHPVYFFKSACPAVTWPDDISDTAFCHCDAKIGMRIVTP
FPSGYCLVGSAPLVSLTDILNRVVKLDTRLASEYPGILEDKGPFDSGIYAKGRCVRVP
HCYKVGPGGELSRLLKIIICHPEESDKSAYLKNAFKVSNLLHHAPGDSVTKNGHLVYA
ITDENEGFLESKTKNNLPKTITDLAEKIERTTEKPLIDWAATAVWPKLHDTIQRFFPD
DRIGQFASVSFMHSGDNIIQVKPQKGNNFFCINHKHRNHTQTVRVFLTLHSTKESEVT
VTFMSQCFAAKCNHNSPTAHFSFMVPITGT"

SEQ ID NO 112 CDS 77578..78906
/note="immediate-early gene product; ORF 57; similar to

Kaposi's sarcoma-associated herpesvirus ORF 57"

SEQ ID NO 113 /translation="MRYVFHALICFIGGISSSDFDDSSSDEMDDLSPTPEPEPSTTPN
SFPEGPKSQVVALPKIRKRSRSETPVKIEHRSPLNRSRSRSRTRSGSGQRSNQSGRYV
KRFKPTVDAPRHREPWHRGGKGKAPFIRRDAMAGRGRRTYGHDYRGKAALTRSIKESI
KKMHLPSTMLSRAHDKKVFEGLLPRHLGQCFQVCLPAPPPLQPEVFTDRQLTAIVKSG
GRRDALVAKKVSLAKLTSLYKPLLTFVTGRNNQAHWLATRKNTLASAGLEALAAFIEE
GLAWAQVCVSQNRSLNDSNLDIILDSSQSVCTWFISKIRHLHIQCFLENQGEVSLVKQ
LTYLVCINNRLAEAANLAGEVKLNFKLGMLIGFALTLPALLAEHKLSGESLYLFRSFL
EKYRPGDVMGLLNSIVVEHYTKCRSAECVITTHAMVGSGENNKGLFFFPV"

SEQ ID NO 114 CDS complement(79266..80513)
/note="R6; similar to Kaposi's sarcoma-associated
herpesvirus vIRF K9"

SEQ ID NO 115 /translation="MATWRPPQSGGPSAMGLREWIVTHANLATYSGLFWADDEKTRVV
LATTTPWSVGFDYLRDGKMYEDYCNQRNIPLPSGRSRLGQAKARLLGAIRKSAYFIEE
KDVLRRSFSFANVVFRLRSDEEMLCRLCPRASGVAAELRGLRFRMFKRKGADEAGRVS
EYTVKQLLGLLRTRPAGTFTMTAPATEASATATASGEDGRQDNSQGGAVALPGEHALP
LSASSGLSACLAPSVDDPWGFMHIQVYYYGVLQAQTFTHSGMGVRLSTRPTDKNEHHV
CMAPGPLQLWLPPAPYMDDDFMLSRLVNALHALEDGIVLCSCQYGIMMNGYGFLNLWF
RGNTSNTSEPRRVPSGVGHRVFDTDEYMLKLAQSPRPSDPGPPDPFAQIWVSAWSLYE
EEDQSQAPICIVVHQREIYRHFE"

SEQ ID NO 116 CDS complement(80686..81933)
/note="R7; similar to Kaposi's sarcoma-associated
herpesvirus vIRF K9"

SEQ ID NO 117 /translation="MAGRGVDIRAWLVAAVESGEYRGLVWENEDKTVVRVPWNKVTAD
RSVWNSEKFFDDYCNMRGICQGEKPSHYGRFRKMRFLYDMRHHKSIRELKFINKAYGR
PGARYRLFRLLPEPVVSCAMCNLMSSTETQCLGLISEFQYDQGGGSGRERRRVFTATV
LARSRMDKNKRVREHRLPGAIQLTFLYFGETVGLERVHAGIRVCSRPCPVLAGHACCF
QDERTLFLPSPGVVDCQFAREDLRVMHKKCEKGVLITLTDTGICVKNLENREMKVLTN
NEEEYKDLPSRQPVQVFDMVDYLRALARSPNPGDEPPRDYAQIALCLSVQSPNPVDAP
IAMRLRYVCETSSVCGTEGCFYPGTTVTSEGRTDCSFQMEDPGEGTSQSHDPAVELGD
SGPDSMDDPDAGTSGEDDGVACS"

SEQ ID NO 118 CDS complement(82262..83317)
/note="R8; similar to Kaposi's sarcoma-associated
herpesvirus vIRF K9"

SEQ ID NO 119 /translation="MERPVRVTKPSSLRGWLVECCETGRHPGMRWVDEERTLIRIPWN
HDRGSRGVEEAEKNIFIDYCRSRGILHAAGRELTAKECKNWLSSAIRHSQTVSDVSTK
DNLSTPYPDRCRIIRLLPITVRSCARCDQASGTTAMLRGLREEAVNKFGPVGAGVRYT
GAVGAGGEQCWMLRIMFYYYGDRVGEVVTESSNGIRVLPLSERRPQGHICAAPIAEQA
LVPEIPGHLAEFQAEALRFLDKDLLRGLAFWADPSGIYIRWLGHSLAFVQGNVESTGA
VAVLSCANACRAFNLVDYMTAMARTSPDGAALPQACVYLYFGGVPTPEGGVQSTVPLI
IQLWHECLWRALSAANV"

SEQ ID NO 120 CDS complement(83491..84252)
/note="R9; vIRF"

SEQ ID NO 121 /translation="MDSGCYACILDENSEGIINYLEQVCGIGLEPGMPLPAPLPTLVP
PTRSAYARAHRLGVPEAPLPHQIVPFWRLRIQVFYFGVLALDHTSQDRRGVRLHPRPV
PHPGHLCFYGTGFTVWFPSPDREKLTAEQITQIKTMLVAYNEGIYVHGNETGVYVDNR
NRETLYAAGNDCNGDIIQREVMFLSKQKIFNFMGFMRKLARSPGPESHAPCNGATLYL
SQQPGAQESPQVPISVVVCQDELVQGHMNPSRWCS"

SEQ ID NO 122 CDS complement(85052..86209)
/note="R10; similar to Kaposi's sarcoma-associated
herpesvirus vIRF K9"

SEQ ID NO 123 /translation="MAAGESRRGPSRYGMALKEWLTFKADSGLYPGLFWADEQKTRLV

LAATPPSFPNYDYQRDGQHYDAYCELRHIPLPSGRSRLCQARGRLLGAVRKSKYFEED
KEFPTDQFPFTALVFRLRSSEEMSCPVCPRVCALRLELRNMRFAMLGRGMLHALSGPS
VSDQERRYREGHQDGHDAQDDDAAYSSGLLRARLMACAAPSAGDPWGHMHIKIYYYGQ
LQAELLTATGQGIRLSSKPTNKAGHHVCVLDGPLQAWFPPIPQTTESSVVQRLEDALK
WLVDGIIFCSTSRGIMFTITGGPNVWFQGNTVEPYSLPHRAYTGMHVWAFDTDRYLLD
MARSPSPRDTGPPAAFVKLWVSGCSLGEERNSSRAPLSIIVYQTEIYRHFE"

SEQ ID NO 124 CDS complement(86355..87527)
/note="R11; similar to Kaposi's sarcoma-associated
herpesvirus vIRF K9"
SEQ ID NO 125 /translation="MAERDMDLKAWFIEAVESKRYPGVEWDDEDKTIIRVPWNRCTDS
RVDEDYNKIFDDFCSARGVCQTGSHAQKFKKIRMLYAVRSHRYLRELTPPSKAGGVSG
ERYRLFQLLPEVTVGCDLCNLIATTSLHSCSMGSCVREDVFERTRRPRAKAPLRVSVY
KRKSKRLQDSSAQPVLGAVEVSFFYFGENVGVQILRAGSGVRICGLPDPKRPGHLCCA
DNPLTCFLPSSQLIPCEFARADLQALQKTCERGLICVMTESGICVKNLEERNMTALTN
YSENYYELRPSQPLQAFDLLHYLRELARSPTPGDVPPRDCAWIFMCPSTQSENTWDAP
IALKLRYVCNDDVSDDVSNGAAGDDSGDEGPSGAGVGASGTTGSTSVSTLAPYGRK"

SEQ ID NO 126 CDS complement(87894..88961)
/note="R12; vIRF"
SEQ ID NO 127 /translation="MAEGRAGSIRVNRPSGLRAWLLDCCDNDKHPGMHWLDEEKTLVR
LPWNHLKGAGGVSDDERNMYLDYCQFKGIRQTGNRRLSVRECKNWLASAIRHSQTVED
VSTEENLSAPAPNRCRVIRLLPIFVRSCPLCNEADATGGMLLDVRNEVTARFRYLGAG
MEYEGAVGGDGEQCWMLRLVVYYYGRLVGNMEVGSPNGVRLLPAPKRPLQGHVCAGIR
PEQALLPHTPQDMFPHQTSMLKWLGKEIIRGLMIYADGSGIYIRYMGHVPAFLLGNGG
SLEPVDIINNARVLRVFSLAQYLSAVSATPPHGTRFPAAYASLHLGGVPTPEGEPCPT
IPLSIQIWHECLWRACGDAAQ"

SEQ ID NO 128 CDS complement(89122..90216)
/note="R13; vIRF"
SEQ ID NO 129 /translation="MTEIEITHNHLRRWIISNLEANTFPEHLCWCDEEKRSFRISWHR
GMSGMQPVVAYCLDRDLECGRQHNVSECRKRLLRVLRENAGFEQDDARATTTRFGGER
FFYLRPAVDPLCYACILDSHSETVLNYLEAACVHGLEPGTPLPPPAPAEADGAARSVY
ARAARLATVAPPHPDQITPFWRLRIRVFYFGSLVAEHTSQDRRGVRLHKRQDPKPGHE
CFYGTAYKMWLPKPQLDGPLTPEQRETVCEIINGCEEGVFLHGNELGMYVDNRTRHTV
RCAGNDAEGNHAQRAVRSSVKSQIFYVMGLLRRLARSPVPGDTVPSNAVTLYLGGRPG
SSKRPQVPVTLVICQDELTHGDIRAARWIL"

SEQ ID NO 130 CDS complement(90462..91544)
/note="unknown; ORF 58; similar to Kaposi's
sarcoma-associated herpesvirus ORF 58"
SEQ ID NO 131 /translation="MGTYTSEASLAWLSFMSGTVSASPFILCFIYHSLYFVEPLISVE
NIIFSWGAVGLHGLLLLFCIFGLPAWLSRQVDVPCTISAFLITAGSMASTLGVDLPWV
HVSIFVGSCLCLLLCVVAANDVVYLCPTIAHRYYELGFLAAFSVYYFLVLKNLFLAPV
FLLPLVAFIVGGVCSLRALRSHPLYEAGLQRRHAIFSLTSRRYITYSIKQALEVCGWD
FYLVTVLIGGAAAGTLSVGLTTPLLLGLVHYFFVFHVGLFCCLGLVFRSNVLALVYVL
AAAVLLTLTHVLGPGTHNLFTRVCVFTVFLLTMFGAIGCELQIIRKKLQRAANSPRIV
LGVCACGNLLMAVVFFSLNKVELGAL"

SEQ ID NO 132 CDS complement(91555..92739)
/note="DNA replication protein; ORF 59; similar to
Kaposi's sarcoma-associated herpesvirus ORF 59"
SEQ ID NO 133 /translation="MPVSFHYGARVDVDALGSISRVYDHIKGIVKKGVIQISGQGRAP
VLSVLSSVGDAGVLGLRLKNALAPLMVYSDMTDEVSFSFRNTSLGNTFTHTREMFGVN
IAEMNVAFYHHGDESDAEGKPQFVRTTIAYGDNHTSTVHKSVVDEPNLPSFHDRLEQA
GTGNRLFLTVKTLTLLLKWLRQQKTRAKQVVTVSLSETLAVATFTVDGVSKIIDFKPD
TPDAKWTCARGRKLDVGVVSSDLTTHVSLESLVAALNACKIPGFFLPGFRWHANEILE

VEGLPLTDSLADVRLGVMLLKVDPTDRNNAVPGNLSEGADPEGVPELPSPPRTPDLDL
KEQCVPIAEDGAEPTDGGAKSLRTSGSRPEKKHGKRKHSSSPSRGKGKTKTPRATFNP
LF"

SEQ ID NO 134 CDS complement(92868..93812)
/note="small ribonucleotide reductase; ORF 60; similar to Kaposi's sarcoma-associated herpesvirus ORF 60"
SEQ ID NO 135 /translation="MFGLSIVTAAMESPDRFLYASDHPGFLALTQETWQNRWFPSQIS
LHEDSDEVRLLSPTDREFYQFLFTFLGMAESLVNFNIEDLVKEFSNHDVTHYYAEQVA
MENIHGKVYANILNLFFGGNRGDLMIYAKKIVEDATLAKKIDWLHSRVRKATTRAEKV
LLFLVIEGIYFISSFYSIGLFRLRGIMRGVCLANDYISRDELLHTRAASLLYNTMISR
DESPSVAYIHGLFREAVEIETLFIRSKSRDVTMVNVGDIEQFLQATADRILKSINIPP
LFGARPPNACPLSYTSAKSVNFFERDNSEYVTSVHNDL"

SEQ ID NO 136 CDS complement(93794..96160)
/note="large ribonucleotide reductase; ORF 61; similar to Kaposi's sarcoma-associated herpesvirus ORF 61"
SEQ ID NO 137 /translation="MNTETSFSAAKSAKPLTLVTDAGTGGCSSSLDPERCAESLVNSL
KATLGWDIEANSLTGLLWHRIMEDRCLVTVRDYLAVFGERLSDEVRAFMSKHEAALDG
LLQDFKQSKAYTNFVNCGYLSAVRFYDTYVLRTQGSSPIFESVAQMFMRVAVFVACQC
IKFPCLRETLRHLVESETELDEMYLVGYAFHYISSQIVCCATPVLRSAGLRGGQLSSC
FILKPSMATEDKTLKALHEEMSPLLASKSGVGIDVSSFAEHKNITSCLKLINAHVGYF
NDNNIRPVGASAYMELWHHQICDFLNAKMPENQERCHNLFQGVCVPELFFRLYETNPD
GQWHLFAPEVAPNLLKLYGAEFEIEYNRLVAAGKHSSSLPLKSMMYALINTVIKTGSP
YVLLKEALNKHHWCETQGSAINCSNLCAEIVQQPEGQASVCNLANISLPKCLRPHRGE
SGVEPGKGDVTFGFELLDDAVEAAVIIVNACILGGTAPTESVRRGQEERSMGIGVQGL
ADVFAELGFGYLDAESAKLDVEIFQAMYFTAVQTSHEIVLLGEGTPFRGWERSRLAQG
VFHWQTWDGVKPSHPPLERWEQLGRSIAQHGIFNSQFLALMPTAGTSQLTGYTEAFYP
FFANIASKVTSKEEILKPNVTFFKRVKPGDLRTVRRYGGDVASFPEPLKDRYKIFLTA
FDYCPIKQLERAGARAPFVDQSQSLNFFLKEEQATRASYIRDLLLTGYRLGLKTMLYY
CRIQKQTKLNALQCLDQVVGDNISSEGAESNCVQKADGERTKVCLACQ"

SEQ ID NO 138 CDS complement(96163..97158)
/note="assembly/DNA maturation protein; ORF 62; similar to Kaposi's sarcoma-associated herpesvirus ORF 62"
SEQ ID NO 139 /translation="MKTRDANVNKLNDSLMRLLPPPHRVSLSRGRDFSKGVRDLLSK
YVVSTTTGVEAIKDGFLSVSPTCQTYGDFLIYSQTMSSQEPRGTYLFSFKQTDTGSSI
DMLFTPTSLARLSRMDADSAPQTNRIACVWYGHESGLLDAIPNFEELLETGSLHQFLA
PVGPLVQTVHSTFVTKVTSALKGNVVAREPVVTHIGLTLPSDMFVDLDDSCPSSLRDE
PLPAHSSIYVCLTYIRVNNRPALGLGFFKSGKGYCEIAAQLRDFYSGVIRTKYIQLQN
DLYINRLAFGVVCRLGSVPSGLQPSFQSLHFKGAALPVLKFTEFVSNPGSWKLFL"

SEQ ID NO 140 CDS 97157..99976
/note="tegument protein; ORF 63; similar to Kaposi's sarcoma-associated herpesvirus ORF 63"
SEQ ID NO 141 /translation="MASSIPAARADNGDENTGGLYKLTDNLLTCTGSLQQLKLLMEFQ
LKPLPTAHLLSMPTVTRFLNTAFKIDNPLVSFIQKHPVFFLMRVARLPEPVITDHQSA
ETSTGILSEVVNVLNTAIRKPHESPAAKDNDYLDNRAILAMITEYIHHVTSRTPSGIP
PTPPMGISHLPCVEQILHETHRQYWNLTLPESLFIDIGEVASPLQTWLILSYCKKLQL
APPPLFPPVDELARRLVTGHHELFVPLSTSLETYITMPVSKQRAFEIYSVFAKSKNIV
DGTPILAFTDTELTTFTPELLFLYDFVIESLCKNQAYGCSRNAIEHFIKKGIDFMAEL
GAFIEKTCGYRSTVSLSNVRAVKARLASCGLSKEACEDFRAMILMTPHETTPKWENFT
DFLEMVNQLTLYGFYFYECLNQYSPTSISLAKIQNILNRVDAEQSDRALWRTPLIGSF
PFPWKLNNVLAFFKPSTPVATLQKIYKAIPSYLMRSLFEIAANKSWGNIALAESAPLT
DIQTAEPDQGPVSAQVIAKYCSRLQISATDYDAAIVSSPGFAAEFIKTKLYPILSEVL
RNTSKKNRSLFQIRWLIVFAAEDARDLAPIRRSLALAYFQIMDILEEKHSPESFYNLL
DYLQETFRCIRQVIPEATCPQEFLQYLFTFQNIPIAASFIQTSMTFVDDLKNGIPGIL

DLVSLGAAFYNMKLLYDSTLDTVEIPTEEGQPIVVSMFVFKSTIRVLEKLLQEAVIAL
TQTSEPMYAAHIRLMQHLTYMQKIAGHEIMTTQLPSVFHEIHEGYLQCFKRFKRLMLH
VTGSCCYSLTRYFGFLYQPPLIPDTIVQKILNFNDKTDTTDDILKSLSQPVRQGPLSA
ENESSSRLSKNNVELLQKLYDDFRTASTNNNPTSIKLEYSGNYNETQVSVDWSTYNLV
TYTAPDDTLKFTPVNTEALDRMFAE"

SEQ ID NO 142 CDS 99980..107626
/note="tegument protein; ORF 64; similar to Kaposi's
sarcoma-associated herpesvirus ORF 64"

SEQ ID NO 143 /translation="MELPPIFSKFKIEGVATTHQADCRFGQYAGSQCLSNCVIYLAQS
YFNRESPVTDTNDLDDVLRQGATLDFILRRSGTLGYNQYAQLHHIPSFIKTNEWTAAI
FQSQEYFGLIGLDAAIREPFIESLKSILTRNYAGTVQYFLFICGDKAGAVIIKNKTFY
LFDPHCVPHVPNSPAHVISSSDPTAILEYVSPPDREYTGSFLYIMPSEYVNPEHYITN
HYRTITFAKVHGPHIDISTGIEPCTIEDIPSPPRSPDVTSKSSNLARVPRTTTDTSSA
KPPPATLSGLRGAEPPTSYPDPATNDADTKLLTPAPAQTAVDHPEFQTTPGATLLLSE
LSASRGRKRKLSSLQRYSDSDEASSDDEGAPRRRVHDDAISAEVIWMDDDISPLYSPS
ATPSFDDVFDSPPMSPEFTYEDATEDTDGAFLEQIARDAETPFSAFDDLITDHDFSSL
DKKIEQLIKYEAPSQHLPNISDKQNGRAVREAAALQAMDKIMINIILEHGLITDAQAR
GPSACKNVLQFFILWGEKLNIPISDAKQVLELDLQLIPLHTAISEGKFKQGAFKKHLT
TKINRCLASMRATHADAQKKLASAFNVEGSQISSSEAKISVRALKEQIANHLSPGFLA
VYSADEVKHLRDKIQDLKTGIEQRNKEIQQEELFFDAMLTALDTFQPPPKTAFPMEIF
PHRKTEVMLDHLASITTRLTEDATEALNNYLETPPDQGTHITNIPNFSSIVANIISTL
KILTYAENDMQLNVTPMATYRRQLLYLGGELATIFNLEWPYETVPPVQELPLVARAKA
KMESVTKMEKNQQALDQILGDAETLLDTITATSGDENPVRAMSIPILETYITNAGALI
GSSRNQRFEKLKAAIHDLASSESFIIMLLNNTRLDNISDNLAKIDGILTNNTRFLSNA
TVSKTLQTLGGSLIRECVEALNKRSPSSLNNARLLAVQTILGHASVPDHETLTRIVSG
VASAQKESAGDDPDRWTRVTGHLNELKLVTTQSRVDKATRRKLLMIITRDLKEAEVSQ
ETVLETRWQENVLKFQPSTSKEIEDFLQSAPSAKARKFAEKHLRTLITQFNGHERPPS
EATAVPMDYTPTPIPTPQAVSTATAEKGKAAWNKIQQAFQDFNFHLIDASDWQEMASE
YSRHGSSLPGTVGPKLVRFMESISNTLDDILTQKLASLLPNGPAFRPPAFDWIAPYQT
RVNAFLKTIGLPMVRNLADKIHHQCQTVSHAVQSADLQQATVGTSLERPAAEYCRILS
DMQVAFNDHGIAVRSEAAAYTDAINSPANVVTPPKPNLEAPKKLITATDALTVEDFPD
FLKTSILQQEQRLIALQRAEFQQLEASISAAERLRQSTRDEIAGKMATAITQLLPRAP
VAISSRPLNLSKPIDFLSSTVYDKILDKEPYETAIAGFAWLEIATKSVMVYSQQNETQ
QLNVLLSEVEKQSTVAQRLHDLELSAKNTDDVKVLKQALDELAPLRVKGGKTTVDAWK
QKLESIESLLRATRTAGEISSELERIGTQAVGTITVRDLGTLSDQCREAANFLRQASL
PEGFSDIGTKLSELQAYIKYKKQFLEHFETTQPNVFQRFPLSQNITENVPARPAMDSV
ARLTNHLHVRGSAPHFTTWIETLPTVDPEKPTHVPAHGGAPLHRQITYSNVLEALFSL
CSTTLTPVPTAPGLEIATRARRGAEAATWMDRQWPDIAQTLQDVLDTYEHTTAHANRD
AAFNTFLAMCVFTQIIRGASRAVTLPKLPSTAVDFPEEIVLTPRECTTLVTAMWPTLA
AAILRLKSYSEALGLMSRFLPLMFQALPHLTLEAQVKNGPHNTPPQLRCFAKTEAIPY
FPAQWQSANLEQSLWGQTDFLQICDNNQRKARVAAVTWALTTIDGVVLDQLWSTFKPM
TAASDDTYVDLVETLHLTTFGPRGPTPRRETTTEHPPYEYGQPTGYCISGQSTTPVQA
SNTPVSAFEAVLGAMVFHVPIRIFLAATPKRLGQARGGMGLLTPILECVPDVEPFKSL
YNAPRKPVPIETLPASLHPHDERQVFLRQAQWLSYRFTPHEAARSSTPPLLVVIDPEN
LVTATYSSGGPANFESRPFYVMPGPYPPDWPKTLSVTSNTSVTHLSHDEICNLFTTLS
REHGTVQGRDIFAAAPTNVTPEQTANPPAWETDNRLITQTETAKKPHIIPASPKARTD
PPVETTTHHSQGQASQHANSNVNQPGQITSHASRNTPSTAPQASSSPEKFNTQTVPRL
ISQTSETAHINQPASGQVTEPKGIFGTYKPRVLTEPAKPANAGVASRQPEATTTVPKL
PINPPTARVFIGTASKLSPAVEESHGATPDAHQSKIDREKYAESRPRRTPHLEEGPRE
PHVNTPTSAHINVPSSQGQKTVHGRENPGLQTATPSAPQPTASNPRIQYTLPRTDGRL
LHDESEVESTPTEEVKRSPKTQDVSHGPEPDDSRWTAPLGPTIEIHRLEHPQILKNIT
SLTVPTPRVTPIPPTNIWIPLSHVNIQHEEITRAKNVLMRFIQNVRRKLQASSDALSE
AIARIKFLYL"

SEQ ID NO 144 CDS complement(107637..108146)

/note="capsid protein; ORF 65; similar to Kaposi's sarcoma-associated herpesvirus ORF 65"

SEQ ID NO 145 /translation="MSSLRVKEPIVQGRLEHDYPNHPLVAEMNNLPQGDMSPAQYAIA
KRNYLVFLTAKHHYDMYMQKKNGILRKDHLRGLRGKKDASSSISGVLSGSGSAAPSVA
PVASTLGSNSFTTISSGPHSLIGSMGPAPGGGGPGSVASSGIGSTSLSPSDATTLDTR
RSSQNKKSK"

SEQ ID NO 146 CDS complement(108152..109498)
/note="unknown; ORF 66; similar to Kaposi's sarcoma-associated herpesvirus ORF 66"

SEQ ID NO 147 /translation="MASGRLPNLAEDEAACHGRGSYPAHRWLDGSRLGLDLAASIRSI
GLCPECYVCFVTYGLGAWDGRPPKWACTLISAPSFQTALNEIATGWRPDNPPKNGDVR
SRLHDIGRSLLEAYAWVLRCICTGVGCPSDEGLSLTAVPRSAWSRYLVVSFQRACCLV
CKTLNCRQRFPLVTCLPQHALDLPVLRKKWNGGGCVSMQLNVPSISRRLGANLNESVP
GPSDAGLLASLRELAPTVPCGNPFNALLRSLTFRALLSMSRVVLPIGESTETEISRDL
GQKVLAYNVLFPCISLPVWSQVVARSVLEKTVPAPRVVVCLECGYCLNFGRGKFETVN
FPPTNVFFSRDQKEKQLSICATTGRVYCSYCGGSHMRVISLFEITCVGDPYLRCVLAN
NAAHAIRDANSLVSVVVPCLASPDCATGLLKHLRVAELFYLTSSISSLSCGKCNRS"

SEQ ID NO 148 CDS complement(109524..110198)
/note="tegument protein; ORF 67; similar to Kaposi's sarcoma-associated herpesvirus ORF 67"

SEQ ID NO 149 /translation="MSSGKRLVDELCDLVVSYLGPSGISLDLERCQDGAPVYAKGGAV
PVCTVRLQHGCVYHLEFVYKFWLHKLERLAYPFAPCFVIINNGLATTLKCFLCKPRDA
DAQFGKNLPINSDVYLERNSSVSLGQDDFMKFKARLVFSGDLNVYSSMVICRTYFTEH
RQVLQFLVVTPKSAKRLKTLLRTVFALTGHSDGLGALRRTGSVARPSGSELKDIGRGE
RAAMTN"

SEQ ID NO 150 CDS 110609..111982
/note="glycoprotein; ORF 68; similar to Kaposi's sarcoma-associated herpesvirus ORF 68"

SEQ ID NO 151 /translation="MFVPWQLETLMRHWPSLRGLVEQSFLPGTPDGAFNSPVLIHTQD
SLQPASSCRVCSLLFTLVRTFPPPDSFFEDYGWLCLTCLYAPRSWTATLMVAADLLEL
THVYFPQCVKDGPVYTAQSILGIDVQLHFFATRCFRPIDREQILHTSHLNFLQTEFIR
GMLEGTIPGSFCFKTSWPRTEKDDQQPTVACCSVGRGSHTNRDNRLPEDLEEAFNSTN
AEEKPSLLGVFSATWAESQLLGSDTQQADTHLQPSAFPTPEDADQSQGPCLMHPTLNL
KTKNHTASICVLCECLAAHPDAGPVLKDLRRDILENMENNVKLVNRISYILNDPDSLS
HVRDEHLRGLIKRCSAQEIHKHFFCDPVCVLNTYSHCPAVLFKCPPPEKYKKLKARLA
TGEFLDCNRIFDCETLQTLAVLFKGSQLAKIGKTTSLEIIRELGFQLRRHNIQITHPF
QTSNLYI"

SEQ ID NO 152 CDS 112004..112897
/note="unknown; ORF 69; similar to Kaposi's sarcoma-associated herpesvirus ORF 69"

SEQ ID NO 153 /translation="MPKQPRSRLASRAPYAPSVRRPDGPQSTRPASRHGSCKSEIMQW
KKLVSDTQFFSALTRRHELGVDFLREMGTPICTSKSVMLPLNLKTIAPGRCVSLSSFG
HSSNMGFNCSSCTPTDRSAVSLDANALGEDSARKNSELCSVALTFYHHAEKVVQHKGF
YLSLLSHSMEVVRKSFTQPGLLYAHLVLKTFGHDPLPIFTVDADERLALWAVFHTRDL
HLGETSLRLIMDNLPNYDITVDCIKQTYIMKFTPSRPDNATVTVPVNSICEAVATLDC
TDEFREEIQRGTAIINSQGLL"

SEQ ID NO 154 CDS complement(119211..119735)
/note="FLIP; ORF 71; similar to Kaposi's sarcoma-associated herpesvirus ORF 71"

SEQ ID NO 155 /translation="MFPHKRLVDFGRHLEADDREAVLWLFDRPASDDTPEGFANGLCP
STGEPGIPLPVLLEAVFLVGRLDLVSTFFLLDVGFIIERLRSSPSYFSPYKHLMLSID
RQLSERDVKNLVFLTGDQLGRRRNQSPTFFRWLSQMEKAALVSPSNYMVLSDLLQAVS

RRDVAKVVAANAPG"

SEQ ID NO 156 CDS complement(119794..120558)
/note="cyclin D homolog; ORF 72; similar to Kaposi's sarcoma-associated herpesvirus ORF 72"

SEQ ID NO 157 /translation="MASVGPVPTGTIDPVLYQDRALSNLLAHEASFVTSTACYGTIQT
EVTVGMRVILGTWMRSVARAHQADASVFPLAVSILDRYLECRSIPRRRFQRLGAACLF
LAGKIRDLNPFKAAFLCFCAAEDFSVADLLKQEKSILKALRWKLEAVLPTDAIGPTLF
KSGFTKEQLFALHSQVVESVHKAIVNPATGGLPPSLVAAACALFSLGAAAPPPARLAE
AVGVSAATLAAAAESVATTLREFDEDHILSNARGSS"

SEQ ID NO 158 CDS complement(120866..122212)
/note="latent nuclear antigen; ORF 73; similar to Kaposi's sarcoma-associated herpesvirus ORF 73"

SEQ ID NO 159 /translation="MWGSRQHRSGIVSGHGLRSSCRGHCGRRGGTREQAGRRGRGRGT
AAPAAAPAPPAPTTSGPQVRAVAEQGHGSDTETATESRHGSSQGSPSGSGSESVIVLG
SPTPSPSGSAPVLASGLSPQNTSGSSPASPASHSPPPSPPSHPGPHSPAPPSSHNPSP
NQQPSSFLQPSHHDSPEPPEPPTSLPPPDSPGPPQSPTPTSSPPPQSPPDSPGPPQSP
TPQQAPSPNTQQAVSHTDHPTGPSRPGPPFPGHTSHSYTVGGWGPPTRAGGVPCLRLR
CTSHNSHEDEAPERQQEQEGEERQQQPARPPRPPRPPRYPIPIPYPSSEEEVPRKYRP
QRRFYRQVLGPRIDPPRPGPWCHGVIFCNSDPYSLYRLARCLQFPGIRASSVRVLPDA
PGSPVIPAFCITVFCQSRGTAKAVKKARRRWERHHPSAPHFQASIVRMDRGLPIQH"

SEQ ID NO 160 CDS 122866..123627
/note="R15; similar to Kaposi's sarcoma-associated herpesvirus K14 and ox-2"

SEQ ID NO 161 /translation="MSGGITLTLLLATLATVRCALQTHYAAVPVHSTASLGCVLTTPH
DVLIVTWQKQESPSPVNVATYSSEAGTVVQPPFAGRVDIPEHKLTRTTLKFFNATLED
EGCYLCIFNAFGVGKLSGTACLTVYVPLSMSVTFYPPINPTQLVCRAEASPAPSVNWT
GVPPELCSEPEVFPRPNGTTLVVGRCNVTSVDPEDLENATCLVTHIGGLAAARPLDPV
FSDPLEGTSHYVVGVVAAAAVLGIFLTGVFLYRSM"

SEQ ID NO 162 CDS 123924..124952
/note="G protein coupled receptor; ORF 74; similar to Kaposi's sarcoma-associated herpesvirus ORF 74"

SEQ ID NO 163 /translation="MDALNNNLNLLMDFLSNYSNSYSSYDDNMSYTLDTESTLCRLTV
VFPPTVYAIICFFIFCITLFGNALVLYIFFKFKALANSVDVLMAGLCCNSLFLCASFL
FSWLLYVAPQMLTSATCKVEIFFFYLYTYFGVYIVVCISLIRCLLVVFSRRPWVKHGA
SGFLCVCVSLIVALALSANASLYRTALRHPETSEWICYEDAGEDTVNWKLRIRTTSAI
CGFLVPFGLMVLFYGLTWCMVKSTKLARKGAVRGVIVTVVVLFLIFCLPYHLCNFFDT
LLRTGFLAETCYLRDVISVAMHICSLLQSMYSAFVPVVYSGLGSLFRRRVRDTWSVFR
CFSTSGSL"

SEQ ID NO 164 CDS complement(125057..128953)
/note="tegument protein; FGARAT; ORF 75; similar to Kaposi's sarcoma-associated herpesvirus ORF 75"

SEQ ID NO 165 /translation="MAQRTNPRWAAAALSPEEEAFIHDNSDAESVLALVPEQCFSEFL
LWLVTRPSDNFDNDDDDPALGVIWHLLAPLVNYAPLETRSAHLQGHHTISLPYGPDLM
RQPTTRSSEIVQCLRDSGLDRTLRLEVGRHLSCQTRRFVADRVPPGTLAALTLGTLVE
YDVRVQRQLPVTLQSTAWRPLPERDPICAAVMLPLQRNILPLAVQASNGNSYTVSRYA
VMARRSYSCVFQRLPCENVTHIADSFTHLHSAIQTGAGALQNILFHATLLPGGEIRSA
LCGFYATTPSVGAFSRARHRAINTTATLHCQQLARTGTPVLGGFLKTVHSATTSEANV
ITTTSLLSCVPQAYTFLRRSLFSQPIICLGSFEPVDGDGNQRSLYLGSAAGITRITQT
LSLAYEILEGPLFTSINRAHEPASVIGHLGALVSRGGLRLFVSQLPPTILSQLTATPD
ISRETVNDILVNKFLNVSACVVFAVLPRDTEPEPGPLDAIRRAARICGCPFAVVGETC
EELGIQFVNDLELWNPGAWPIRQPTSAEVIATFGFDEQPVSSNWLVRPEEPEDGGEQA
PSPTDWGLFRLASVVDQLLRCPTVGSKEFVTRHVDRCSNGLVAQQCEVGPLGRPLSDY

HIVNHTSVFTDRMARVPIYRPQPITRQDATERLVSPETWVTQGRGRNRWVGQCVAYGE
QAYKMGINAAVGARYAICEAVTNIMLAHVRRLSDITLTASVGWNPEDDQAWLLQHTLF
ACKELCRDLSINFAITSAGSTPCLSEELISATQQHQTVAPVPFNAVVITATAEVKSSR
QRVTPDLKATGNLIVLVSFPVPHLTQGSTFEHLCLLPSPTLPDVQATHLANLFMLTEA
LLSRGLVVSGHDVSDGGMVVTAIEMALAGNRGLQIRIPSEETPLQWLVSETPGVIFEI
QPQHVDEVRQACQNFDCRATVCGTVGQEGLSERIVISHNNDEVYSQTLTSVAANWTSF
SDEQWYSWGPSFTPAQELYRKDYGCNQHNLGHLAEVCRNSELTLFATPSRPPAVAALV
TPGAPLPRALMAAFTNVGFDVAAVSTDDLRGGNILRGFSGLTIGGTVGIEDSYVGARC
AIMGLLNDPGCYGGLMAFFRRADTFSLCCGEFGFQLLGALGLLRETPHDTPGPKTPDQ
WDIHLEENASGNHECLWLNLHIPQTTISIMFRVLRGLVLPGWANGRYLGVRYPRDAME
YHLNQQQRIALNYHTGNADPRMFAQHYPRNPSANSAVAAITSPDGRHLASLVDPAVTF
HPWQWAYVPPELADMTVSPWALAFQSLFLWCIRNRQ"

SEQ ID NO 166
PCR primer
CCTATGGGCTCCATGAGC

SEQ ID NO 167
PCR primer
ATCGTCAATCAGGCTGCG

SEQ ID NO 168
PCR primer
ATATTAAACACTCGCCGC

SEQ ID NO 169
PCR primer
ATGAGGGGCCTTTTCGTGTGC

SEQ ID NO 170
PCR primer
CTGAATCCCGCTGCCAAGGCC

SEQ ID NO 171
PCR primer
ATGTTCCCTGTCTGGTTCGTC

SEQ ID NO 172
PCR primer
TTACATCATAGCTATTGCGCG

SEQ ID NO 173
<213> Macaca mulatta rhadinovirus 17577

nucleotides complement(23398..23668)

SEQ ID NO 174
<213> Macaca mulatta rhadinovirus 17577

nucleotides 25065..25368

SEQ ID NO 175
<213> Macaca mulatta rhadinovirus 17577

nucleotides 25518..26525

SEQ ID NO 176
<213> Macaca mulatta rhadinovirus 17577

nucleotides 114979..115383

SEQ ID NO 177
<213> Macaca mulatta rhadinovirus 17577
nucleotides 115385..116413

SEQ ID NO 178
<213> Macaca mulatta rhadinovirus 17577

nucleotides 131731..131926

SEQ ID NO 179
<213> Macaca mulatta rhadinovirus 17577
nucleotides 132333..133719
partial terminal repeat

SEQUENCE LISTING

<110> Oregon Health Sciences University

<120> Cloning of Rhadinovirus Genome and Methods for its Use

<130> 53683

<140> PCT/US99/26260
<141> 1999-11-05

<150> 60/107,507
<151> 1998-11-06

<150> 60/109,409
<151> 1998-11-20

<160> 179

<170> PatentIn Ver. 2.1

<210> 1
<211> 133719
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 1
gatcgggaaa acgcgagggg agcgggggac aggggacggc gtgtgcgtgc ttgtgagaca 60
ccgggtacgg ctgcctgcct gctcgctggc ctgcttgctg aggggacagt aggcctgctt 120
gctcgctggc ctgcttgctg aggggacagt aggcctgctt gctgagggga cagtaggcct 180
gcttgctcgc tggcctgctt gctgagggga cagtagggct gctggcttgc tagtagggct 240
gctcgctggc ctgcttgctt gctcgctggc ctgcttgctt gctcgctggc ctgcttgctt 300
gctcgctggc ctgcttgctt gctcgctggc ctgcttgctt gctcgctggc ctgcttgctt 360
gctcgctggc ctgcttgctt gctcgctggc ctgcttgctt gctcgctggc ctgcttgctg 420
aggggacagt agggctgctt gcttgctaag gggacggtac gcctgcctga tggcttgata 480
gtagggctgc tgggctgcta gtagggctgc tgggctgcta gtagggctgc tgggctgcta 540
gtagggctgc tgggctgcta gtagggctcc tgggctgcta gtagggctgc tgggctgcta 600
gtagggctcc tgggctgcta gtagggctgc tgggctgcta gtagggctgc tgggctgcta 660
gtagggctgc tgggctgcta gtagggctgc ctgctggctt gcttgcttgc ttgctagtgg 720
ggccgcttgc ctgctactag ggctgctgtg cagctgggag aacagagtag ggctgccggc 780
cagctgcgtg cgagggcgtc cgagggccag acgaggacac gggacccggg cctctccccc 840
aggcacaaag cagtagggct ggccagggga aacagtaagg ctgattgctt gctgaaaaac 900
agtagggctg ctggtttgtt gctaacggaa aaaggggagg tgtgtgtatg cttttgtgac 960
tttccagcgg agggggtaca gtgcacggcc aagttacaag cacctgctta acttgctttg 1020
gctctgtgcg gttttgttgc taggtactag tagtaacaca tagtatttca tcgcatggcg 1080
cctgcatacg ctccacagca tccgaaacac gttatttcta tagactaact ttagtgttct 1140
tcttggttgg taccatttta cgaagtttgc tccgttgaat aataagaaat ccgttgtggt 1200
tacaatacac ctgccaacgt attggatggt tccttttgcc atcaaccttt gcttgcattc 1260
taacattgaa ttttcacatt acaccctagt cttctcatgc aattaggtat gtctttgctc 1320
gtgtttaatt ttctctacac atttctgtgg aaatgtttgt gttggtttta tttatgttat 1380
tgcaacctgt atcggtggag cttttgcctg caaaattaac ttctgttcca acgtggtgtc 1440
caccacatcc cggagatacc tacttgctaa cctgccgcgg gacgtctacg gccagagacc 1500
agcgaagcac acaatggttt cgcaacaaca cgcttatgcg tgggagtaat ttctacggca 1560
gactggtatc tgtgactccc aatgctacga tatctgaccg gtatgcgtgt caaacaaaaa 1620
caacaacgcg gagtaacaac atcgattttc gggtaagctc atcgcgcctc acgctccaag 1680
aacggtgctc ttcatacggc tatacttacg cgaataacac aagggtattg aggtgttact 1740
ctggtggaaa cgtaacttta agaaacgttg tctttcattt aaacggtaca gcggtcatca 1800
acggtactac aacaaacata catacatttg tgttaacaga aaagacagga gggacgtatt 1860
tctgttctgc gtttattggg aatgaaaaat tctattctca gacaattaat gtgtttttta 1920
cttcatttac ctttaaacct acaaacgaca ttcccaatga gtcacatttt aataaaactg 1980
ggcaaataca acaaacagct agtgtacaac atcctgaaaa ctacgttgtg ttctctgttc 2040

```
ccgtttttc tattggcgtt ttaacaggta ttgcaatatc gttgattatg tgttggttat 2100
ttacaatacg ctgcaacgag aactctgaat catcaactaa tagttatgca agccagacaa 2160
gctacattca accctcccat aatcagcgtt ccaatactaa tgaatgtagt cgccatacct 2220
acagaaatgc tcatcaagaa gagagtattg aagaactacc aaaccaacac acaagtgaaa 2280
ctgattcttg ctgtcaatta gttttacttg aagtgaaaaa tgtagcctac gatggaccgc 2340
aggaaaacac aattaacgaa gttatggaac agtatgatga tgtggttgta aaaaatatag 2400
aacaaacatc atatgaggat aatgttgagc acatggacta tagtgatact ataaatccca 2460
attttaatta ctacagtgga ctaatattgg aagaagtaga tgaagttttt tacaatgaac 2520
tagaaaatca atatcatgga ttaatactgg agaatttaga tcacaatgag tacaatcatt 2580
taaatgaatt aaacatgata gaacaatatg attggttaga ataaataact tgtgggttat 2640
ttttaaatta aaacataaac aacattaaag cagcattttt gtgtaaatcc tttatttatt 2700
aaaatttttt tcatatacct gaaacttata tttaattcca ttttcctcaa agtttgtatc 2760
tttgcccggt atctctgata acatggtgta ttcggtaaaa ttaattgatg gaaaaaatac 2820
atcacaatca aaactctcca taatgcgcgt aatatatagt tttaagggac acttatagtt 2880
aagaacactc tcataaacag atttcctcc aataacccag accgtgttta attgttcttt 2940
tagtttgtat tgtctataga aattaaacgc atcgtctagc gttctcgcta gaaagtgtgc 3000
tccgtgcggt ggttcacgta attctctgct caaaataata ttaattctgt tcaccagagg 3060
gcgcttcttt tcagggatgg aaaaccatgt tcttttcccc ataataacca cattcttttc 3120
acctacaacc gatggcgtag acgtcatttt ctgaaaatac atcatttcgt ttctaaggta 3180
cggccaaggc attgttccgt ttttaccaat tcctaattgt tcatcaactg caacgatgca 3240
gttaactgta atgtccatgg ttttagttgc cactgacggg tttaacacag aagtatttca 3300
caattataaa caaataaccc acgtgacatg tacttactaa tgtaagtacg taacgtgata 3360
tagttaatca tattcgcttt cacataatgg acaaaacgaa aaaatgttac ctgtgttaag 3420
cctaccttgt ttggatggat ttatggttac aatgaaaaat aaaaatatat atataacgtt 3480
ttttacgtaa aacatattgt aatttaatca cgtgatgttt atgttggcta ctaacacaca 3540
atgtttgcat aaaagcactt atgttagtga gcttcaatct ttatatttta ttcaaattgt 3600
ttaaagaggc atgtgaatat atttaacacg cattcagaaa taggcgcaac gctgtgccgt 3660
ataactggta aaaacatgac gtttaaactt tttcctctgt ttttattaca cgccataatg 3720
tacgtccact gcgatgaaaa ctgtaaacct ccacatttca cggaatatcg cgtcaagtct 3780
aacacagaaa aggacttata tagtgttgga gaaacagctg aattaatttg tcgtcctggt 3840
tatgttacaa atacaaaaat aataacaaca gaatgtttac aaaatggtac gtggtcaaca 3900
ccaaattttc catgcgacag aaaaagatgt cccacacctg ctgacttgct gaatggagcc 3960
gtgcacattc acgggggaga taatgcctta aaatttggat ccaatatttc ctatgagtgt 4020
aatgaaggtt atgatttaat tggtagtaat gttcgttttt gtattttaca agacacagaa 4080
aacgtaaatt gggattcaaa tgaaccagtc tgtgaaattc agaaatgtat taaaccaccg 4140
gcagtggaac atggggacta cctacctaac caagatgttt ataactatgg agatgcaatt 4200
acatttaaat gttcattgtc gtatacactc gttggatcaa caacattagt atgcacgtca 4260
aacaaaaagt ggtcaaactc tttcccaacg tgtttaatgc tcgtatgtga aagtccacaa 4320
atagacaatg ggtacataga cattggcttg tccagaagat acaaccatgg acaatcaatt 4380
actgtaaagt gtagcgacgg gtacaacatt gttgggcctg aaacattaac gtgcacaaac 4440
acaacttggg ttccaccatt acctaagtgt gtattagtta caaataaccc aagcacaccc 4500
atgccagaaa cacccatgcc agaaacaccc acgccagatt atcaaaaaat aaatttgtca 4560
accgctaaaa ctgcaacaac accaaatgcg tttgttacaa ctgttgtttc tccagagaaa 4620
gacgacgtta cttgtgtaaa gcctcatttt gagcgattca tggtaaaggc tgaaaatgac 4680
aaggaaaaat acagtgttgg tgcgagcgtt gagctaatat gtcgaccagg atttactaaa 4740
atgcagtcta cagtttctgt tgaatgtttg tccaacggaa catggactgc tccaaatgcc 4800
aagtgtcata gaaaaaaatg tccaacccct caagaacttt taaacggaga gtatatagtt 4860
acaagcggag aagatgcttt taagtacgga acaaatataa catataaatg taatgaaggt 4920
tatcaacttt taggaagtat ggtgcggatt tgtatgctta aagacgattt aaaaacagtt 4980
gactgggagc caaaagcgcc tatatgtgat attgaaaaat gtaagccacc gccacaaatt 5040
acaaacggaa aataccatcc ggtgaaagac ttttatcagt atttggacac cgtaacattt 5100
tcgtgcaatc gtgacttttc tttagttgga gatgaaatga caacgtgtat aagtaatacg 5160
tggaataaac cgtttccaag atgtgaacaa atcacttgca gcgctcctaa tattgcacac 5220
ggaaagctgc taacaggttc ttcaagcgtt tacaaatacg gtcaatctgt taccattggt 5280
tgtgaaactg gatttactct aattggcagt gaaatttcta catgcaagga ttcatcgtgg 5340
gatccaccac ttcctacgtg cgtgccagct gtttcaatgc cttctgacac acctaaacca 5400
gaaaccaaaa aaccaaacac gccaacgcca gaagcaccca aaccaaacac cccaaacgtt 5460
ggaacacata caccattcaa accaccacca caaaatccac caatagcacc cccaatgagt 5520
aaatggaaaa ggcatgtcgt gttagttctt tttgcaagtg tcgcgtcctt gttattcgta 5580
cttgctgccc tttattgttg ttttctaaaa taactgtttt ttgtcttcag caggttcgcc 5640
aggcaaactc gcacgcatta accaatctgc caaccgccgt tgattccgga attaagttta 5700
```

cattattcaa ggttgccaat aaaggtggtt taaaaatatt ctattggtgt tcattgtttt 5760
atgttgaccc gtttatagtt atcgcgccac cttgtggcta cattatatag cacgatcact 5820
ttccacgtta tactttcacg tactatgact catacgcctt aacgtcacgt ggcgtgcgat 5880
tgtggccggg gctgaaaata acacaagggg tacataatcc atccaggcgg cacacattag 5940
acacggttta taaaactata tcggatgcgc caacaatcac tgtcgctagc gacactgata 6000
gaaaaacatt ttaacgtttg tttagcgaac ttgaataaca cataatggct tccaaaggca 6060
acgccggaca acccctggaa gataatcagg ggtctcgtgc cccgataggt gcgtgcggat 6120
acgtgtacgc gtattcgaaa caagactttc cctttgccga ggcgtccata ctcggcaaca 6180
gaccatctgg atctggcgtt ttctcgctac caatccttta cggacttaca gttgaacacg 6240
aattccctct caccgtaaaa gccgcataca aaaaagttga caccacgacg ctcgccgtta 6300
aggtgacgtg ctttcacaga gaggttattg tgtttcacaa tgcaagttta ttcaggccgg 6360
tgtttgacgg taccggtctt aacgaactat gcgaggaagc cagggctctc tttgggtaca 6420
cgcagtttat agaaccgggt ccacctcaca gcatatggaa ccctctggaa tgtccgcagt 6480
taccggacaa ggatgagatg tttcttggcg ttgttgttac ggaagggttt aaggaaagac 6540
tgtggagggg ctgtctcgtt cccgcggtgt tccagaccca gcaggtgcag attgccggac 6600
gccaggcgtt taaagtgccg ttgtacgacg aagacctgtt tgcacctcac ggtcatagaa 6660
tgccaaggtt ttaccataaa gacgttagcg cgtacctcta cgactccctc tttaccagca 6720
tcgcccaggc cctgagactc aaagacgtga cggcggtcat ccacgccaca gaaaagcaat 6780
tcatgcagga ccattacaaa attgccaaga tagtgcaggc aaaacagttt tcaacgacgc 6840
tgccgaaaac gacagacggg tcgtcccaca tgattgtgga cagcgtcgtc gccgagctcg 6900
cccttagtta cggctgtatg tttctcgagt gtccccagga cgcgtgcgag ttgctgaact 6960
acgatagctg gcccatattt gatggttgtg actcaccaga ggctagggtt aacgcgttag 7020
agcgctggtc ggccgaacag gccgttcacg tggcgggtca actgttcgct gccaattcgg 7080
tgctgtacct aactaaagtg cagaagcaag cgcccagggg acaaaaggga gacgtaaacg 7140
tgtacaactc ctttttcctc caacacggac tggggttttt aaatgaggcc acgatcaagg 7200
aaaacggcag cgaagccttt aagggcgtac cctcaaacgc cctcgatggt tcttcgttca 7260
cgccgtatca cctggcctac gccgcgtctt tctcgcccca tctgctggcg aagttatgtt 7320
attacatgca gttcttgcaa caccacaaaa gctccacgaa ccaggcgttt aacatggtcc 7380
attatgtcgg caccgccgcc aactcagaga tgtgcacgct atgtcacggc aacacgccgg 7440
caacgtgcct caacacgctg ttctatagac tgaaggatag gtttcccgcc gtaaccaccc 7500
ctcagcgcag ggacccctac gtggtgaccg gaacagccgg gacctttaac gacctggaga 7560
ttctgggcaa cttcgcgagc tttagagacc gcgaagagga cggaaacccg gccgacgagc 7620
acccaaagta cacgtactgg cagctatgtc agaccgtgac agaaaagcta tccgcgattg 7680
gaatcaccga agaccacgat aatcacgtga acctcatcac caacatccaa agttttctca 7740
gggtgttcaa gggtatcgac tcaattgtgg acggagaggt catgaagttc gttaattcga 7800
tgattaaaaa taactttaat ttccgcgagc acgtcaaatc ggtccatcac atactccagt 7860
tctgctgcaa cgtgtattgg caggcgccgt gcgcggtgtt cctgaatctg tactacaaat 7920
ccctgctgtg gatcattcag gatatctgcc tgccgtactg catgatctac gaacaagata 7980
atccggcgat gggcatcctc ccctccgagt ggctaaagat gcattttcag acgttgtgga 8040
cgaactttaa agcggcgtgt ctcgaccgcg gggtcctcac ggggtgcgaa ctgaaaatcg 8100
tacaccggga catgttctgc gacttcttcg acaccgacgc ggggtccaac ggcttaatgg 8160
ccccctttaa aatgcaggtt agaatagccc gagccatgat ggtcgttccg aaatcaatta 8220
aaataaaaaa tagaatcatt ttttccaaca ccgcgggatc cgaggcggtg cagtcggggt 8280
tcgtcaaacc gacgggaacc agggacactt acgtggtggc cggaccgtac atgaagtttc 8340
tcaactcgct gcatcgcgcg ctgttccccg acaccaagac cgccgcgctg tacctgtggc 8400
acaagatctc ccagaccaac aaaaccccag ttctgaaaga cgtcccggac gacgagctgg 8460
cggagctggt gtcgtacgta aagaccaaca gcctcgcgtt cgaggaaacg aacgtgctgg 8520
acgtggttcc ggattcactc atgtcgtacg cgaggatcaa actgaacggg gccattctaa 8580
gggcatgtgg ccagattcag ttctacgcca cgacgctgca ctgcctcacg ccggtgctac 8640
agacgatcga tgccgaggaa tacccccacg tgctgggctc cgcggcaatc gccacaccgg 8700
tggcttacct ggcagaaata cgcggccgca ccgccctcac cgtccagacg acggcgcgtc 8760
agccggtcgc cgccacaggg cgcctgcgtc ccgtgataac cgttcccatg gtagtcaaca 8820
aatacacggg ggtcaacggg aacaacaacg ttttccactg cggaaacctg ggtacttcg 8880
cggggcgcgg cgtggaccgc aacctgtggc cggaaagctc cccctttaag aaaacgggcg 8940
tcagcgccat gctaagaaag agacacgtca tgatgacccc cattatcgac cgcctaataa 9000
agcgagccgc gggacagaca atcagcacgt tcgaggcgga aagcgttaaa aggagcgtgc 9060
aggcgctgtt agaggataag gacaacccta acctattgaa gtcggtaatc ttggagctta 9120
tacgacacct ggggaagggc tgccaggact taagctccga ggacgtgcaa tattacctcg 9180
gtgactattg tatgttgacg gacgaggttt tatttacgtt ggataatata gcacagtcag 9240
gcgtgccgtg gactatcgag gacgcgggtg ccctaataga ggatcgccag gacgcagacg 9300
atcttcagtt cgtagacagc gacgatatcg ccaccgcttc ctgtcagccc cccgaggaac 9360

agctaccgac ccctagcgcc ggcgccctac tggccgggaa gaagcgaaaa attaacgcgc 9420
tgctgagcga tctagacctt taggaaaacc gtggcaggcg ggcaacaatg gccagggaac 9480
tcgcagcatt atacgcgcag ctgtcggccc tcgccgtcga cttgagtctg gttatctttg 9540
cggacccgcg aagtatcgac ggtgcccgca ttctaaaaac aaaaacacag atagagaacc 9600
tgaaccgcga ccttctgccg ctgctacgcg agcaaaactc ggtagagacg tccagcctgt 9660
cgctcgaagt ggagcacctg gccaaaaaca tcgaggacaa actcggcgag ctggagcgca 9720
gtctgcggca gagatattcg agccgagagc attttgaaac actacacctg agacccgaat 9780
gtcactatca ctctacggtt acttttcagt tttacggggg cgggttaata gatgtaaaca 9840
tgtgcctaat aaacgatgta gaactgctgt gtaaaagact agggagtgtg ttttattgca 9900
tcggtgcaaa cgaagctctg tccggattga accgggttct gacgtttctg tcaacactgc 9960
ggggtatctc cccgatcccg cacccagacc tatacgtcac gtcagtgcct tgcgtacagt 10020
gcctgaggga aatcgaactc gtaccaaatc aggggtccag tttactcgcg gtgttggcag 10080
accgacactg cgatcacctc tgtaagaagg ttagggcgga gccaatacac ggcctgtttg 10140
agacagaact gagccagctg ggtctaaaag taacaaaacg ttcggacgcc acgcagcacg 10200
gcgtccggtc ctctgcagat cagttaaggg agtcgtcgct ggcggccata caagatcaca 10260
atatattcaa acgggtgtcc gcgtcaatca tggaactatc caatctaatt tattggaacg 10320
ccgggcaaac cggcctccag accgggaccg aaaacgagtg ctcacaaatg gccagactgc 10380
taacacacga ggccgatatg cacgagcacc gtgcgctaat aacacccaaa ctaagcgcga 10440
ctcacttcta cgactgtttc cgaccggatc ccatagaatc cctgttctgc ggcggtcttt 10500
ttaactctat agacgacacc ataaacgcac tgagccggga ttgctccgtg acgttctttc 10560
aacaggcaaa ctataccaac gttatgcgaa aacaaaacga gctgttcacc agactcaata 10620
gcatcctgcg tcaggggagc gcgggatcgc aaaaaccggc cacccccctcg gagccacgga 10680
ccaccaccgt ggcggcaacc gcggcaagcg acgtcatcaa agacgcacag tatcgcaaag 10740
aacagtacat gaaaaaggtg gccagggacg gctttaaaaa actaacagag tgtctgcaga 10800
cgcagagcgc ggtgttggca aacgcactct gcatgcgcgt atggggggc gtcgcctacg 10860
gcgaggcgtc cgagctggtg aaccatttc tcctcaggcg gcgcttcgtc gcgcttccct 10920
gggaggcgcg ctgccgctcg gatcagattt tattcgaaaa ctcaaagtac attaaaaact 10980
cactatattc ccagcgcctc agtcgcgaac acgtagagat tatcacgctg cagttttacg 11040
gcctgataac cggccccctg acgcgccaga gcgatctctt tcccggcccc gccaacgtcg 11100
cgctggccca gtgtttcgag gcggccggaa tgcttccgca tcacaagatg ctggtgtcag 11160
agatgatatg gccccagatt caaccgaaag actggataga ccagacattt aatcgttttt 11220
accaacttcc cgaaggtgat ctcaacgcgg tacaaaagtc cgcctggtgc tttatacgag 11280
agctcgtcct ctcggtggcg ctttataatc gcacgtggga aaagacgctg cggatatttt 11340
ccctagcgcg cgagaaactc tccatctcca acctagacgt taaaggcctg acgtccggcc 11400
tgtatctaac gtacgagcaa gacgcgccgc tcgttctaat ttctcaaaat accggctgga 11460
tatttaaaga cctgtacgct cttctgtacc atcacctgca actgtccgac ggccatgatg 11520
ataactaacc gaacgcgtcg tctcctgcgg gcgtgggtcg tgataatcgc gatcggcacg 11580
gcggttggcg aaaacgtcac caccccaag ggcgcgacca ccaccgcgaa gccaacgccg 11640
ggcccgtcga cgcccacacc tcccgagaac ccacctaggg ccgaggcgtt taagtttcgc 11700
gtgtgcagcg cctcggccac cggcgaactc ttcaggttta acctggaaaa aacgtgtccg 11760
ggcaccgagg acaagacgca ccaagaaggc atcctgatgg tgtttaaaaa aaatattgtc 11820
ccgcacatct ttaaggtcag acggtaccgc aaggtggcca cctcggtgac cgtctatcga 11880
gggtggaccg agaccgccgt gaccggcaag caagaggtca tccgaccggt gccgcagtac 11940
gagatcaacc acatggacac gacctaccag tgtttcagct ccatgcgcgt aaacgtcaac 12000
ggcatagtaa acacctacac ggacagggac ttcactaacc agaccgtgtt tctgcaaccg 12060
gtcgaggggc tcacggataa catccagcga tacttcagtc agccggtgct gtacacgaca 12120
ccgggatggt ttccgggaat ttacagggtc agaaccacgg tcaactgcga gatcgtggac 12180
atgatcgcgc gttcggcgga accgtactcg tattttgtca ccgccctggg agacacggta 12240
gaggtgtccc cgttctgcca caacgactca acgtgctcgg tcgcggagaa aaccgaaaac 12300
ggcctcggcg cccgcgtgct cacaaattac accatcgtcg acttcgcgac ccgccagccc 12360
accaccgaaa cgcgggtctt cgccgactcg ggagaataca ccgtatcgtg gaaggcggag 12420
gaccccaagt cggcggtctg cgcgctgacg ctctggaaaa ccttccccag ggcgatacag 12480
acgacgcacg aggccagcta ccacttcgtg gccaacgacg tgacggcgac cttcacgtcc 12540
ccgctctccc aggtaactaa cttcacgggc acgtacccct gcctcaatga tgttattcag 12600
aaaaccctca acgccaccat caagaagctg tccgataccc acgcaacaaa cggatcggag 12660
cagtactacg aaaccgaggg gggtctgttt ctcctgtggc agccgttaac gccgctaagc 12720
ctagctgacg agatgcgcga attaaacggc accacgccag caccccccac cacaacctca 12780
accgccaacc gcgttcgaag aagcgtcggt acgaacgagc aggcaacgga cgacctagcg 12840
gcgccccagc tgcagttcgc ctacgacaag ctccgcgcga gcatcaacaa ggtgctggag 12900
gagctctcca gggcgtggtg ccgagaacag gtgagggaca cctacatgtg gtacgaactg 12960
agcaagatta accccaccag cgtaatgacg gcgatatacg ggcggccggt gtcggccaag 13020

ttcgtgggcg acgccatctc cgtgacggac tgcgtggcgg tggaccaggc gtccgtcagc 13080
atccacaaga gcctccgcac gtccaccccg ggatctgct actcgcgccc cccggtcacg 13140
ttcaggttcc tcaacagcac cacgctgttc aagggccagc tgggacccag aaacgagatc 13200
atactgacgg acaaccaggt ggaggcgtgc aaagagacgt gcgaacacta cttcatagcg 13260
agcaacgtaa cctactacta caaagactac gtcttcgtga aaaaaattaa cacctccgag 13320
atatccaccc tcggtacgtt catcgccctg aacctgtcgt ttatagagaa catagatttc 13380
agggtcatcg agctgtacag ccgcgcggag aaaaagctgt ccgggagcgt tttcgatata 13440
gaaaccatgt tcagggaata caactactac acgcaacgcc tggcgggact ccgggaggac 13500
ctggacaaca cgatcgacct gaaccgcgac cgcctggccc gcgacctgtc cgagatagtc 13560
gcggacctgg gcgatgtcgg ccgcacggtc gttaacgtgg ccagtagcgt gataaccctg 13620
ttcggatcaa tcgtgagcgg gttcattaac tttataaaga gtccgttcgg gggcatgctc 13680
atgatcctgg tgattgtggc ggtcgtcctg atcgtgtttg cgctaaaccg gcgcaccaac 13740
gccatcgccc aggcccccat caggatgatc taccccgaca tagacaaaat gcagccctct 13800
ggcggtaaag tcgaccagga gcagattaaa aacattctcg ccggcatgca ccagctacag 13860
caggaagagc gtaggcggtt agacgaacag cagaggtcag cgccctcgct tttccggcgc 13920
gcgtcagacg gactaaaacg tcgctttagg ggatataaac cgctggaaaa cgaagaggct 13980
caagagtatg aaatgagcaa ataaccacac ccacacgcct gtacttgccg cccgccagag 14040
ccgcgcggcc aatcgattcg cgcacccggc cggtccccga caccctctag cgccccccggg 14100
cgtccgcggt ctgtatcaat catggatttc tttaacccgt acctgggccc tcgcggacca 14160
cgccccccact cacacagagg caccgatgct cccgcccctg ccggcgccgg agccgttcag 14220
ccgccaccag acgtttgcag gctcatcccc gcctgcctcc gaacgccagg ggcaggcggg 14280
atgatcccgg tcacgatccc gttcccgcca acgtacttcg agaacggtgc tcgcggagac 14340
gtgctgctcg ccaacgaacg gtccatgtgg acggcgcgcg accgcaagcc cgtcgccccg 14400
gacccccaag accaatccat cacgtttcac gcgtacgacg tcgttgaaac aacgtacgcg 14460
gcggacaggt gtgccgaggt acctagccgc ttccaaacgg acattatccc aagcggaacc 14520
gtgctcaagc tcctggggcg aaccgaggac ggcaccagcg tgtgcgtgaa cgtgttccgt 14580
caacaggtat atttctacgc gaaggttcca gccggcatta acgtcaccca catcctccag 14640
caggccctca agaacacagc cggccgtgcc gcgtgcggct tctcgaccag aagagtgaac 14700
aaaagaattc tcaaaacgta cgacgtcgcg gagcatcccg tcacggaaat cacgctatcg 14760
tccggttcca tgctctcgac cctcagcgac cgcctcgtcg cgtgcgggtg cgaggtgttc 14820
gagtcaaacg tggacgccgt tcgccggttc gttctggatc acgggtttac cacgttcggg 14880
tggtactcgt gcgcgcgcgc cacgccccgc ctggcggcca gagatgccag gacggccctg 14940
gagtttgact gcagctggga ggacctcagc gttcaagcgg accgcagcga ctggcccccg 15000
taccgcatcg tggcctttga tatcgagtgc actggagagg cgggatttcc gtgcgccacg 15060
cgcgacggcg acgcggtgat ccagatctcc tgcgtcttct acacgaccag ggaaggcgcg 15120
cccaatccgc caaacatact gttcagcgtc gggacgtgcg accccatccc ggacaccgac 15180
gttttggagt ttccgtcgga atatgacatg ctggtgtcgt tcttcgccat gatccgcgac 15240
ttcgaggtgg actttttaac cggctataac atctcaaact tcgatctccc gtacctaatc 15300
acgcgagcgt cccaggtgta caaccttcga ttaaacgaat acacaaaaat aaaaaccggc 15360
tccatctttg aagttcacga gccccgtggc gggggagggg ggttcatgag gtcggtctca 15420
aaaattaaaa tagcgggcat cgtccccata gacatgtacc aggtgtgtcg cgaaaagctc 15480
agcctctccg actacaaact ggacacggtg gccaggcagt gtctgggtgg gaaaaaagag 15540
gacgtatcgt acaaggacat tccccctctg tttcgctcag gtccgggcgg cagggctaag 15600
gtgggcagct attgcgtgat ggactcggtc ctggtgatgg acctcttaaa aatgtttatg 15660
atacacgtgg agatttcgga gatagccaag ctggccaaga ttcaggccag gcgcgtcctg 15720
acggacggcc aacagctccg cgtgttctcc tgcctgctgg aggccgcggc cagggagaac 15780
tttatcctcc cggttccaac gcccgaggga caggggggct atcagggcgc gacggtgatc 15840
aaccccattc cggggtttta cgacgagccg gtcctggtgg tcgattttgc cagcctgtac 15900
ccgagcatca tccaggcgca caacctgtgc tactccacca tgatacacgg acgagacctg 15960
cacctgcacc ccaacctgac gccggacgac tacgagacgt tcgtgctgag cggcggaccg 16020
gtacattttg taaaaaaaca caagcgggag tctctgctgg gaagactgct aaccgtgtgg 16080
ttagaaaagc gaagggcgat ccggcgcacc ctggcggcgt gcgatgaccc gtcgctaaaa 16140
accatcttag ataaacaaca gctggccatc aaggtgacat gtaacgcggt ttacgggttc 16200
accggggtgg ccagcggcct cctcccatgc attaacatag cggaaaccgt gacgctccgg 16260
gggcgcacga tgctggagat gtcaaagtct tacgtggagg ccctgacgac ggaagacctg 16320
cgaacgcgtc tcggtcgcga ggtgaccgcc cgtcacggcg cgcggtttcg cgtcgtctac 16380
ggtgacaccg actccctctt tatcgcgtgc gacggttatt ccgcggaagc cgtttccgct 16440
ttctgtgacg atctggccgc caggatcact gcggacctgt tccccccacc cattaagcta 16500
gaggcggaaa agacgttcaa gtgtctgctg ctgctgacga aaaagcgcta catcggggtc 16560
ctattgaacg acaaaatggt catgaaaggg gtcgacctca ttcgcaaaac ggcctgcaag 16620
tttgtccagg agcgatgccg cgccatcctg gacctggtgc tccacgatcc ggaggtcaag 16680

gctgcggcgc ggctgttgtg caagcggccg ccgcacgcgg tatacgagga ggggctgccg 16740
gctggcttta taaaaatcgt agaggtcctc aacgcgagct atctggacct ccgaaacagc 16800
gtcgtgccca tcgagcagtt aacgttctcc accgagctca gccgccccgt ctgcgattac 16860
aagaccacca acctgcccca cctggcggtg taccaaaagc tggcgagcag gtgcgaggag 16920
ctgccccagg tgcacgatag aatcccctac gtgttcgttg acgcgcccgg gtccctaaag 16980
tcggacctgg ccgaacaccc ggattacgtc agacagcacc agattcccgt cgcggtcgac 17040
ctatatttcg acaaactggt gcacggcgcg gccaacatcc tccagtgtct gttcggcaac 17100
aacgcggaca ccacggtggc catcctctac aattttctca acgtcccgta taagctgttc 17160
tcgtgaacgc caattggaga acgccaacat aagacgccgc gccagcggag tccgcaggga 17220
gagctcgagc ggcgaggagc gaccaacgga gaccgccacc atgctggtta acgaactgtc 17280
ggtggtcctc ggcgactggg aggtgacttt tcaccggggt agattcagct tcgtcaacct 17340
cacccgcctg caaacgttca agggccacgg gggctacgcc agggtccgac tccccttctc 17400
gctcgaccag ttactccacc aacatttcgc gttcggactc gtgacgcgtc tcaaggaact 17460
gcccccccttc tccgactgcg tggcccttat cgccccgttg gattccggcg gcgacgcgga 17520
cgcggcgcgc gtggcccccg ggttcgtgct ggactcctct cgcccgctga ccgtgtgggt 17580
aaacgcgagc gggcggcaca cgatccggtt ctgcctcctc tttctaaagc cgatcgacct 17640
ggagcgcgcg gtcacgtacg tcttcggcga gaacggcggc gcgcgctcgg agggcacccc 17700
aaagcccacc tgcgcgaccg aaagcctgcc cggtggaccc ctgcgcgtct ccggcgaggc 17760
gtctcagacg tcgccccatt ctttcgttgc gtattttccc acggccaact cggtggcctg 17820
cctaagcctg ttgcggttac aggtgaggcc gttttcggat gacgcggcgc acagggacgc 17880
gcggatctcc ccgaaatacg tcacgtttag taactccggg ggtaacgtct gcaaggcgtc 17940
cgttcacacg ctgtccccgt cgcggtgtaa aacggcgcaa atggaaatca tctacgctcc 18000
cggggacccc aacgccgaga tagtcctggg ccagtccgga cccgtcctgc ccacccacac 18060
cggcggccgc gtattggggg tctacgccga cgccgaaaaa accatccaac ctggaagctc 18120
cgcggaagtc cgggttcagt taatcttcca acagggagcg gccgctcggg gcgatctggc 18180
gtttctggtc acgggcgtgg caccggagcc cctattcgtc gtcaccccgg cactcttgct 18240
ttccggttgc acaacccacc tgcgcctatt caaccccaac ggtaccccca cgactataaa 18300
aagagacacc cttgtggccg ccgccgcgcc ctgccccgtg gtgcgattaa gctccgccga 18360
cgacgcgccg cgagacctcg tcgcgtcacc agacaccggg gcgctctcca ttaacgcgtt 18420
cacaatcccg gtcggtttcc caggggtggt ctcggcggag tgtcacgtgt ccctacgcga 18480
caacggggtc cacgaacgca tgaaccattg acggcaacga tgggaacacc ggtgcgtttc 18540
tttcgcggcg agtggcagac ctcgagtcta gtggacaacg gcacgccacg gtacagctcc 18600
ctggtgtggg ccgccactat tcacgacggc tacctgacac tggtgaacag gtcagagctg 18660
tgcgtcacgg agaggtctcc gtgtctgccg gcatgcccca gcatcgggag actggtcggg 18720
aagaggtttc ccggcttcgc ctttgccagc gccactctgg gcgatcgggg aacacgcacc 18780
gtgttctacg cgttcggtca ccgcgacaac ccactggaca tagtacccgc cgtggtcgag 18840
cgcgcggatc gcgagctggt gctgcgggtt cacgctccgc aaacaacgcg ggtgtcgcga 18900
tacggactta aggtattcgt ggcgatcgtt acggtggtgc gcccgcccgg ggtgttccta 18960
cactttccac aagaccgcgt tccgatcgcg ctgacagacg cgtgcagcca ggagggctcc 19020
aggctaacct ctgaagagcc gtggataaaa attcaaggct ttcccgtcct atctgacgag 19080
accgcgcacc catttctctt aacccagaag accaagccct ttaccgagcg aaagttttgc 19140
cgcctgatca tggacaacga ccagcgcagc gccgtcaaca ccgtctacct gggaaagcag 19200
cacgtgaggg tgaccgtgac ccgccccccg gaaacaatcg tcaccgacgg ccccgtgacg 19260
gcgaccctgt ccctcaccgg taatgcgcca atcgcctttc gccacaaccc atactttgaa 19320
ctcccgtggt cgtccacaac ggcgatattc acgcccgtgg tgtacgtggg cctgaccgtg 19380
tgcatcccac ccaactgtag caaattcgta aggtacggta acacctacgt ctcggcattt 19440
aaccgcaagc tgacggcgat tattagcaat cacgcccaca acggcgggtt ccggattcag 19500
gactgcgagt ggccaccgaa ccgggagata gagattttgg taaccaacgt gtcccaggcc 19560
ccggtgtaca tcagcaccgg gacgcagctg gggcaagcca tcttcgtgtt cgcgccgcgg 19620
ttcggtggcc cggcgaaact gcggcagctc ctcggccacc gatcgcgcgc cctggagctg 19680
ccgggcgggg tgacagtgga cagccaaaaa ctgtgtaggt ttgagaccat gtacctgttt 19740
tccacgtaaa ttactaataa accgtttgct cgtatcgctc acacaacgcc aaaccgtctc 19800
tcattctcgg ggtcgcgcgc ctcgcgaaca cacaaggtgg ctcaaacacc cccctcccgc 19860
accctcgcca cacaaaacca gttaacgcct tccgttagat gcagtttatt tattatttta 19920
ttacatcata gctattgcgc ggcgcccgtc ccgcaaaaac atctgtagat attccagtat 19980
gcgaaacgcg ctgagaacaa cgtcccgggg gctctgccct cccaacgcac gcacggtttt 20040
ttcattagac tccgcgccac ctatcttgtt atttacggga agctcctcaa ttagagagtc 20100
gagggcggac agcaccacgg tcaccgcgga gccaatggcc gcggcgtccg cgtccccggg 20160
cgcgtcgtcc agcagcatcc tcagtcgcgt caggtaactc tcgtaggccc ggagcccgcg 20220
aaccatggcg ttcatacact cggccccgcg aaactcgcgg cgctggcact caacgttaga 20280
catcaatata gggaagtgaa caaacgacat gatgcgcggg tactcacgca cccgacagag 20340

agtggagtgg tggcacaggt aaacgaacca gcgctgtatg tttaacccca cttcagacag 20400
gcgccccctg gtcctgtcaa gagaggcgcg gttgccggcc cactggggga gaacgttaat 20460
tccagcggca gtcgggggag gcgccagcgt agggctggcc gcccaacacg acaggtaaaa 20520
caagacgaac cagacaggga acatgaccgg ttaaattact cgcttacaat cgcgggcggc 20580
ggccggtcaa cgccaggtcc attaaaaaca caccggccgc caaccccaa cgcggggccg 20640
cgccctggaa cgcggttcct tccaatcgca aagaaccgcg tcacaaaaag ggctcgtttt 20700
gaacccattt tgtgccatcg ggtttcgttt tcagatacgg aaacggctcg tccaaaaaca 20760
cccaacgggg gtgtgctcaa acggccatct ccatctctat gtggggatgg gggtcgtagc 20820
cctcgagact cagatccgcg cgcgtaaagt cctccagacg cgccaccttt ctcaaaatct 20880
tcagccgcgg aaacggacgc ggggtcctcc gcagctgaag cagcagggga tcaacgtggt 20940
tgttgtaaac gtgggcgtca cccaaggtgt gcacaaagtc tcccggggtc aggcccgtga 21000
cgtgagcgat cagatacgtc aggagggcgt agctggcgat gttaaacggg accccgaggc 21060
ccatgtcggc ggacctctgg tacagctggc aggacagctc cccccgagcc acgtaaaact 21120
gacacaaaac gtgacaagga gggagagcca tccgcgcgag gtccgcgggg ttccacgcgc 21180
acataacgat gcgccgatcg tggggccgcc tgttaattag atccaccacg taacgcagct 21240
ggtccacccc ctgaccctcg tagttggcgt cggcccccct gtactccgcc ccaaaatgtc 21300
tccactggaa cccgtacacc ggccccagat cgccctcgcg gcggtccccg aagccctgcg 21360
ccgccaaaaa ggcgcgggac ccgtgcgcgt cccaaatttt tacgccgcgg cgcgacagtt 21420
cggtggagtc ggtggagccc ctgataaacc acagcaactc ctccacgacg cccctccaaa 21480
acacccttt ggtggttaac agaggaaact cgtccctgag gttatatcgg gcctgaagcc 21540
cgaacacgga cctggtgccc acgcccgtcc tgtcctccct ctgcacgccg tgtttaataa 21600
ttaaatccag gtgcgccaag tactgcagct cgccgtggtc gccgcggcac gcgcagcgga 21660
cggccgcggc accttcgggt gcggaaaaaa cccggagtcg tgcggccgcg attccagcca 21720
cgcaaacggg tataattttt ttaacgtagc agataccgag atgcaccagg acaatcataa 21780
ttacaaccga ggtgctttac aactataacc gtaaacggct gaagacgaaa cttattttaa 21840
aggcagcttg ggcggggccg accacaggcg cggcccggct gttagcgagt aataaaccca 21900
cgtggtctcc ccaagcgtct ccccgagcgt gagcgtctcc ccaaacgcca aaaacaacac 21960
agataataaa ataaataaca tgtttatttt ttataaactt aacgcgcgat tatgcttctt 22020
aacaggggcc aatagaagca acgtgcgagg cggcattcgc gcgctcaaag gcaaccgcaa 22080
caccggcgct cgtaacagcg taaacaaaca actattagcc attcgtaacc gtaaaatcca 22140
accctctgcg tcccgagcca ttcgcccagc tgcgtctacg tcgagggtgt ttctaagtcg 22200
cattttgagc ccttggcacg gccgcgtctg cgttaatggt gtttctaaat cccgctgcca 22260
aggccctctt cggaatcgtc aatcaggctg cgccgggttc ttctgccctt tttgggacgc 22320
ttgtccaccg cctgaagcag tttcttcact gcgtcgtccc cgggattggc acacagcttt 22380
ttaccgcggc gagtgtttaa tatcacggcg tccaccgagc actgcgacga ggtgtgggag 22440
taagagacca ctaaaccggg tggcggcaga tgagttacat accccaaaca gcagagttcg 22500
ggcgcgggtc cgctcatgga gcccatagga aaggcataat caactacaca cgcgaacacc 22560
gcaaaaaaaa cgcacacgaa aaggcccctc atggtgacaa aaacacaaaa caaagacggg 22620
gtcgctcgtg cgttgccacc agcgccacgc tcgccgtgct cggccacggt tgacggcgtc 22680
cattaatacc cgcggcccgg gcccgcctcg acggaaacca ggcgtggtta gccacctgac 22740
gcacgtgcct cccgtccaga tgtggcattg cgagtcccaa aacgggcgtg tctaactacg 22800
gtcgccccgg ggttcaaaga acggggatcc cccaaacagg taaaaagctt tttgtccaaa 22860
cagaaagcca acagcaaccg aaaggaatct ccaaaacaag caagggtccc gtgccggcga 22920
taaacggaaa atttccctca cctaaacaca cgcgccgcgt aaactgttaa aacacgcaat 22980
acttcctagg cctgtgttta ataaaacaca caagggttgt aatttcgcgg ggccctttgg 23040
gccccgcgcg tctgggggtg cgttgtctcg ttgtggttcg caatcaccac cccgcctgcc 23100
cgaaaacagg gcaacaaccc cctggcctag tttttaaaa acttaacacc ggcaagggga 23160
gagggagaag gggtgcggct aaatgggctg ttagcagcca accgcagcgg ggaaaaaagg 23220
gggcaaacgc cgggttaaca caaacagcaa caggcgccaa acccaatata acacaacgca 23280
tattccgtgt gtttcgcggg ccggggtcgg tacggctaac ccgcgttggg cgtttgcaca 23340
agcacgcccc cgccctgcac ccggctcccc gaggtcaccg aagggtcagg aaacatagtt 23400
ttcacggtcg ccagtcgcgc gcgtggcaac atctttcggt gccctccgaa cgacccggaa 23460
aaatccccgc cgcgccgggt cgttcggagg gcaccgaaag atatccccaa acgcaaccta 23520
aagcatcatg tttggggttt cggtgacgcg cgccgcagag gaggccggtg gtggcgctgg 23580
cgaagatagt ggccgcagag cacgggtttc ggtagtgacg cgcctggacg gctgccacct 23640
ggcagcaggc caaaaacacg ggcgccgcga gcggccggca ggtggcgccg cgctttcccc 23700
cacgccgcga tttggcaaaa tggcggaccg gcagcgccgg ttggacgcgg cggacgtcaa 23760
tcacgacgga gagaagcgca gttggtcagt tgtcccgtca atcaccgagg gccgcgccgg 23820
ttggcggata atatgaaacc ggcgcggcca ttggacgctg gcggcagcca atgggcgcgt 23880
tgggttttgt ttacaagttc cctatatata ttatattata ccttccctgt taagggaatc 23940
ccatgttctt aaatctaatc agctgggtta tataattaga atatcagtaa ttactattat 24000

tactgtatta tttattaaat tatttatttt attatttatt taataaatta attcttgcgt 24060
tattatgtaa tcagatactg ggaaatctac tcccagttta ataaaaataa ataattaatt 24120
ctattatggt gggggttcta aatattgtac aaaacaacaa tattttttaa atttattcac 24180
taaaaatgct cccattttac ttgtgacgta catgttagga atgtgggagt ggtttgtaac 24240
tatgttttca aacaacaccc ttttgggccg tctgaaactc atataagcgg aaggctcggc 24300
gttaattgcc actggccgct aaccagtcca ccttgccagt tggagtttta ttgctgctta 24360
tggctttacc tcgcattccg ggcacaagct gatttacgta gcacgcattg gggtttaac 24420
ccggttgaat ttagtggttt gcacgtgtaa aaaagcccac cgaggcttgg gtttaagcgt 24480
ttaaagcgtt gctggtttta taacatgcga tttaattgct atgggagcta gcttgtattc 24540
cgcatgcttt cggtttagta gttgtacaaa acagaattgt agcccgcaaa ttgttactgg 24600
tacccaaaac ggcaagcccg cttaacacgc cccgcgctgc ggcttttgtt cctagcccgc 24660
tcgcaataca atcgagattt taatttaaaa tggggcaatg gcgtagcgca aatttaagcg 24720
cgtcaaaaat ttaaaatgga ctaacgtaac cggtgctttt agtgacctat ggcgaatttt 24780
aaaatttaaa gccgtgggtt ttaacacaga gctgccagct tgtaatacgt tgcatataca 24840
gtgccgttgt gattttatgg ttaagtggat tttaacattg aattttgtaa tggtgtacac 24900
gtggattttt aaattgaatt taaaattgta attacgtgcg cgtggatttt aaaattgtat 24960
ttatgtaaca ttgttaaatt ttaactgcgt tatggttgta ttttggttaa tttatgtaat 25020
tttgctttta gatttgttcg ggtgtgaatt ccagacaggt aagagcctag ctcctaatgt 25080
ttgccttgcc gcctagctcc taatgtttgc cttgccgcct agctcctaat gtttgccttg 25140
ccgcctagct cctaatgttt gccttgccgc ctagctccta atgtttgcct tgccgcctag 25200
ctcctaatgt ttgccttgcc gcctagctcc taatgtttgc cttgccgcct agctcctaat 25260
gtttgccttg ccgcctagct cctaatgttt gccttgccgc ctagctccta atgtttgcct 25320
tgccgcctag ctcctaatgt ttgccttgcc gcctagctcc taatgtttaa caacattaat 25380
gtttaagcac actaaaattt aaagacgttt gtgttggttt ttatgaccag cttggtacaa 25440
aacctgctgg tgattttttta cccaacaaat aataaataaa aaagttaaaa cttatttctg 25500
ttgtctgtgg taattgggtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg 25560
gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt 25620
ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt 25680
cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc 25740
ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg 25800
agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg 25860
gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt 25920
ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt 25980
cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc 26040
ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg 26100
agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg 26160
gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt 26220
ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt 26280
cccgggctcc ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc 26340
ccgttccccg agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg 26400
agggtcccgg gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg 26460
gctccccgtt ccccgagggt cccgggctcc ccgttccccg agggtcccgg gctccccgtt 26520
ccccgattca acgttaaagt ggtttattta gcaaacacac gttttttgcc gcggtttggg 26580
ttgtacatac acgcgggtgc gcattggccc gcccatgcaa aaaaataacg ttgggtttcc 26640
gaaacgcgta ggtgctttag tgtttcatgg cacgttttcg tgtcacggta ggagtggtta 26700
gtttttggt agcgaatgca accaggtagc tcgcctgacc gggaaaggga acgttttaaa 26760
ccgcagtgtt accgcggtgc ggcatacggt ttaacgtgca actgccggtc ccgcgcgtgg 26820
taccttaagt ttacgtgtag caattatggc ggctgttcag ggccctccgc cgccccctga 26880
ggaagaaaat gaaaattctt tgccagttga cgtttatgct atagagggca tctttcttta 26940
ctgtgggctc gggcaggcgg agtacttgca ccatcccgtg tttagcccta ttaaggaatt 27000
tatcagcgcc tttctcaagg acagcgctcg cctgtacgag aggctcttgc gccacaccga 27060
ttaccgctct ctgcggggac taaacgccat aggccaaggg atgctgcaaa taaacacgga 27120
cggacgccac aactggggtc gcgctttggc cgtgttgggt cttggtgcgt atgtggtgga 27180
taaggttaaa gacgacgagc gtcttttaac gttcgccata gccgttctac ccgtgtacgc 27240
gtacgaggcg ctggagtctc agtggtttcg ttcacacggc gaatgggagg gactcaggaa 27300
ttactgcgag cgaatactga ggcatcgccg caacgcgagg agacacatgt gctacggagt 27360
tgcggctggt cttctggcgc tagtggcgct gtttgccatc aggcgatagg tgtgcgttta 27420
accgcgtatt cccccaccct gaccttaagt gagccgattg tatatgagac caataaaaca 27480
aaacaagcta acgtggtatt cgttggaaca ttttttattt gagcagttcc tcgcagaaca 27540
tttttgtat ctgtgacacg ggggccggct gtgcgctggc ctccaccggg gcccgcgcgg 27600
actgtccgca ttctggatcg aggggcgggc acgcgccagc gcggggggca gaggctgaag 27660

gaatggcgtt tgacattaca gattcttggg cggcgggggc tggagttgcc ggctgcgtgg 27720
cgagaccggc gtggtggctc gccggctgcg ggggcggtgg ttgcgtttgc tggatgccct 27780
gcgcgctctg gtacggctgt ggggcgtgga gatgggtgtc cgttccggaa taccactgcg 27840
gctggtacgc taggtacggc gcagacgcgt aggtctgtgg ccacgagaac tgaacgggcg 27900
gtggttgcgg tcgttggtcg gcgtcctgtg cgccggaggt ctgcgacgcg gctcgtttta 27960
ggtccctcag atccgcctga atgtccaata tgtttttgga cagcgccatg acgtccttgt 28020
gtatgccgac ctcctcgcct ggaaacacgg gctcgtcgga aagctccacg tcccgtttgc 28080
gcttctggtg gcgagacggt gccggagcgt agccgtacgt cggagaatag tacgtggccg 28140
catacggaga gggttggggg gcaaacgccg actggaaagg gaacggtggc gcgacgggcg 28200
gtgcgtacga tgggacggtc agctcaggcg gcggtatcca cgatggatac gccgccgttg 28260
ccggcatcga gggtggcgca gaaacgtacg tgcggtggcc ctggttgcgc atgttatcga 28320
gactgctctg caccatggtc agaaaggtgc tcttgggaat ggtgatgttg tcctccggga 28380
gggcgctcat ggttcgggtg tcgcggttac cgcagtgctg agcgcaggga aactggcttg 28440
cctttaaata cgtacttctc gcgactctgg ccacgccttt atctgttttg agcaggtctg 28500
tgcgatctcg gataaatccg gcgtcgatcg ccttggccat cagtgtctca agaggggccg 28560
cgaattctgg cggtgtcacc tgacgttcgc aaagatctaa acaattgacc gttattttac 28620
cggcttcctc tcgcgtaaga gaatcaaact tggagacgac ccatgtaggg tctggtccgt 28680
aaacggcgat tgacccccta cgcctcccca gtgcgcacag agagacgtgc tgaaacgcgg 28740
gtgggtgatt ggggtcttgc agctcttctg ggtggaggga cgagagggag agctccggga 28800
gccatgtgtg taacatctcc aacaggggt cccgaggcat ggggtttagc tggatctggg 28860
ccgctcgcga gtcccccgcg agccgtgacg ccagggccaa gaaatcgtgg ctcgttagct 28920
tgcccaggca aaatatcccg tgtgttactt ggaagagacc gattgtgtga cctacgtggg 28980
cttcggggac gtgctctatg tttatcggta gagggtccgt atacgggagc agggtcgcca 29040
cgattgaggg ctctaaatac agctcctttt ctatctttgg taggctgacc acgtccacgt 29100
atccccccaac gtacacggga gtcatgttgg gtcggtgtct gcgtgagccg aaggagatga 29160
gcgccggtct tcgggggctg atgtggcgcg ttatccgttg tgagaattta aatacgtttc 29220
tccctggcga gctcaggttt cttcacctgg tgctgtgcga aatgtacaac tacggtctta 29280
acgtgtacct tctgaaggag gcaattgcca ataccggaac gagggacgat attgtcctcg 29340
ggcgtaaggt tccggtggag ttctggaaaa tcgtttacga tgggttaaag gagatggggg 29400
tttctgacgc aacgttgctc tcagaaacca aaagggggc tctgtggctt tacttcaacg 29460
ggcgtccgtg cctgttaaaa ggtctaggcg actacgtttt ttgccaactg gggttaagcc 29520
acagcgtacg cgtggttccg gaaaacctaa cggatgggaa ttatctctat aacctaggca 29580
gcgtcatacc atgtcgcctg cttgtggcac tgagctactg tttggcgttt tgggggcacg 29640
ctgaccacga gccctgggtc aggttgttcg ccggcaagat ttttatactg tacttaataa 29700
tttccggcca cattatgcca cggaagtcaa tcttagaaca ggtgggaacg tccgggtacg 29760
gtggcttcgt cgaggcagtg tgtcgcgatg ttcgcgcggt acacgggatt ccggcgtggg 29820
attttgcgtc cgcggctccg gcgttaacgt ctcaacagac ggactacttg tttgcgttca 29880
ataacagcgt agtgtagggt cgcctcacac taccgacgtg gcgaccggca ggtatcccag 29940
gaccaaaaaa taaactctgt cgtagtcgtc gtaaactccg aattggcttt tcattatctc 30000
gtgtgccaga ccggggggcg atgggtgcga aaggatggcg gccaggcctt cctcggccgt 30060
taaacgtagc gacgtttgag cctgaatcag gctgccagcg tccctgaacg tgagcttttg 30120
gttcaggatg tcaaagattt cgttaaatct ggaaccggtc aggttgacga acggacgagt 30180
gtgttttaga tgggatatgg cctttgcctc cagcgccgct agatacgcac ccgggaacaa 30240
aaacgacgtg gggcgtctg gcgcgcgcga aagaattggg gagatgtaat tttctgtcag 30300
gaaatcaaaa agttgtccct ttttaaggta ttgtttggac ccggacgggt cttctacctt 30360
ggaaaatctc ccacccagga gcgaggccag gttgaggcgc cggtggggcg atctgggtgt 30420
gatttgggcc agctccaata gcacgtaaac caaaactaaa cagcctagtg ccgagaggcc 30480
ggctcggagg ttccattggt acgcttgaaa gggaggcacg tcttcgctga acaaagtggt 30540
cgcaaggatg tatacgagat ctggtccagg tggctgaccg gtaccggggt tcttggaacc 30600
gggtagcagt ccgctctgtg ccaacgtcgc gtatatcacg tggtttgtaa agaagtccgc 30660
gttgtacttt gtttctttgt ttaggggcgc tagagaggcg cgttgcctcg gattcgaata 30720
tgtaaactcg aaacccaggc ccttcgtgtt ttgttttagg tgagagacaa gcgtcgccag 30780
ctttgtgccg aggttatccc agaggccgtt tcgagttggc ggcaacgggg tacctgtccg 30840
gtgacagtag tacccgttga ggaggcacag cgcggcctgc gtgtccgata cgttgcggac 30900
atccgcgtaa aagtcttgtg tgatcgaggc ggctgttgag attacggcag acattagctt 30960
tagagacgtg gagttttgaa agtttgttgt gcctcgcaaa tctctgggga acaccctgcg 31020
ttgcatggcg ttatctgtga ggctactgta ccacggcccg tattgtggca gccattggtt 31080
ttggttcatg tacaggcttg gcaggagctc cagacggaga tcctggttta ccgtgaatcc 31140
cgggtctccg gggctattg tgacaacgaa cgcgtctcca tgtgtctgcg gggcgatttt 31200
gggttgcggg tttagtgatg tttgggcatt ggcggtttcc ctgcccgggg accgcagatc 31260
cccgaccatg ttttgaattg tttctaggtc gcgcatgacc acttctgttt cccgttggtg 31320

cgtttgcagc acgttgtcta gttctagccg gagcaggttt tttcgcagcg tgtgaaccct 31380
tagcgccagc tgttgatttc gcatcacaaa cgcgtttctt ctttcttcta tcagagatgg 31440
tttgtgaata aacagaacgt tgtttggatg gggggcccaa aaagttgcgg taatttgccg 31500
cgctggtttt cgtgggtatc tcatacagca tttttctgac gtgcgcatct tcccattggc 31560
taaaaaaccc gttaagcgcc gccgacgtgg tgttgataaa tgtaggcgga aacgctgatg 31620
actcagtatg aattgtttta agtgttttct ggttcttaaa tagcagccag ggccttaccg 31680
accagtttc gtgtcctctt ggcgcggctc tgcctatata tgccacggaa tctgtcagtt 31740
ggtgtaagcc ctgccatcgc tgtgatattt ttattgcgtt gccgggaaat acagccgcgc 31800
ttaggcacgt cttgagttca attacgtagc acgtttttg atgttcgcta ttaacggtaa 31860
gcacacaatc tgctattcgc ctacctaggg aaacctcgaa aaaaatccta aatttaacgt 31920
ttttggttgc cctgagtgtt gggcgaattg aaagaaactt taaaatatca tcaagcgacc 31980
ggtgtttgat taaccgtcgg tatactgcta agtgggcgcg agttccagcc cgtttgcgtg 32040
aggccggaag ggcttcgagg gtcgcgacgt gcttacattg atttgcaaac gccatggccg 32100
aaggagggtc tggctttggc gacgagcttg tccgtcagat gcgggatcgg aagcctagat 32160
gggacgagtc ttccgacgat actgatgacg ttgacaccga aagcaccgat cttgagtatg 32220
atgatgtgtt ccccgtggta gatacgcacg gcttaatgag ccctggaagt caaaactatg 32280
acgtacccac gtctccgtcc ggaacgccgt gggaattact gcacccggac gccctgtatg 32340
cacatccgag gtgcccgcct aaaagggcgg tggtgccggg cggtggcgcg cgtcccaagg 32400
tgtccgcgtt ctcggctaga cttcaatatg ttggacgaca gagttttggg gatagagaga 32460
cgcggcagct taccggggct cagttttctt ccgagagcga gcacgaatat gcggaaatac 32520
cggagcgtac tactacacgc ccggttgaaa gcggtgacaa gagaaatttc acctccgggc 32580
gtaggggtgc aatctcggga ccctcgtcga caaaacctag tcacggtgcg gggttaacac 32640
gtaagactaa aacgtcgctg agtgttagtc ttaaaaactt gctgcggata aaggacgatg 32700
atgctaaggt tgatgtaccc aggccggtta ccgtgccggt tcatctcatg cagccgcatc 32760
ctatgacgga atatagaaat gcctttttaa tctatctaga gggggtcatg ggtgttggaa 32820
aaacaacgct attgaattca atgactggca tggtgccgca ggaaaacgtc ttaagctgtc 32880
ccgagcccat gaaattttgg acgtgtgttt attcaaattg ccttaaagaa cagcgcagca 32940
tagttaagca aggcacccac gggaaattga tcacttctgc tcgcgtatac gcgtgtcaga 33000
gcaagtttgc gctaccgttt cgtgcgacag ccgccggcat cggtcgcaac ctgcaaccgt 33060
ggctggtggg aaacggtagc acaaagccag cgaattggat tgtttttgac agacacctac 33120
tgtccgcaac ggtggttttt ccgctggttc acgtaaagta caacaggcta acgccggatc 33180
atctgtttca aatcctatct cttttctcgg cacacgacgg cgatgtggtc gttttgctaa 33240
cgctcaacag ttcggaggcg cacaggcgca ttcaaagtcg aggccgtaag gaagaaaaag 33300
gaatcacgca aaactacttg cgacaggtag cgtgggcgta ccatgccgtg ttctgtacgt 33360
gggtgatgat gcaatatctc acaccggagc aaatggttca gctgtgtgta caaactgtgt 33420
ccatagagga catatgcaat atgaattcca gattgactca tcggtttcta accctaacaa 33480
aactacatga acagagtatg atacccatgg tggcagaaat gttagtatcg gttaaagaac 33540
acgtgacctt aatggaggtc tgtttgggac tctttaaaga gctacgaaag cttcaaattt 33600
taattgttga cgcaggagaa catctagatg atgcgtgtgg cctctgggga aatatttatg 33660
ggcaggtgat gtcaaatgag gctattaaac cacgagcagt gaactggcca gctcttgaaa 33720
gctacattca aacgctaacc aaattggaag gcaatggcgc gtattagttt tatatttttt 33780
tttactataa taaggtgttc tgtaacggac aaatacgtgt atgatgaaaa atcaaatgta 33840
gagttggaat ttaatggaac aatttaccaa attaattgga gaaacgtaag caaagaatta 33900
acatctattg taatggaaga tgcgtggtac gattctttac tgctggaacc attgtctgtg 33960
actttggaaa aaagaaaaag tttgctgcga tcaagtatag ttaatgtgca taataatgac 34020
tatacatttt gtaaatcttc aagtgaccat gtaattaacc ttacggtgga ttttaattac 34080
agttctttgc caggttttac tgggaatttt aatgtaatga cacatgcttt aacacaaggc 34140
gttttattaa cgaaacgtga attatttacc aactcaacaa acattatgga tctttttac 34200
gcggaaaaaa taaatgcaga aatgttcaag attacttttg attattccaa tgtaataatt 34260
tccggcatta taacagaaaa ttggatatta gtatcagtta caaattcaag tgtaaaatca 34320
aatatgcaat gtgttgcact gttgtttggt gttcctagca catttccagc gttaaaaggc 34380
tatgtatctt atcgtgattt attggtcgtt aaaaattcaa actatgctct gggtgtaatt 34440
gcgccgaaat cttataatac tcttgatttg gcatttttac caaaaaattt tacagaaatg 34500
tttgtttcag ttatagatag tccgcttaat gccattgatt atctcaaagg aaaacttttg 34560
gctattgaag caaagggtgc ctgtcaaaat ccttcaaatg aaaatgatat cttatctttc 34620
ttttttgaag ttactgcagt aaatttttta tttataaaaa atcttcaaaa gcagcaatta 34680
gttaacgttg ggtgtgttgt acggcatgtt gcagcgttag aaagtttaat gcatttatta 34740
aggctctgct acccaacgtt taagttatat gagctaaatt tggaaacgtt gtcgcatatt 34800
gcagagtccc aggtttttaa cttgccggca aatagcatgt tatcgttatc agtaaatgat 34860
caagaggtcg tattttcaat gtttaaaatt gtatataata cacctaaggt tggtggtaaa 34920
attctaaacg aaattgtata tattactaat tacatgtaca ccaaatacag tgaaaattat 34980

cagctaacca acacatttag gcggaatgta atgaacatgt atgaagtttt aacaacaatt 35040
aaattaaacg tcacagattc ttctgtgttt tatccgtata tattgttcac atccatgtgt 35100
aataatgtag agatatcata tatgattaat caaatagcaa agcccgatga catcaccata 35160
tttcgcgtgt tttcgccttg ttttttaagc ttaaggtttg atttagatga aaacaagcta 35220
cgatccgacg caccgcagac gtcaaaaaga accggctcgg aactggcaca aggagcgtct 35280
ggtttttggc ggctgttgca cgcttttcac gccacgcgta taaacgagtt ttcagtcata 35340
aattgtacca ggctggcatg gaagcaagta accgcgctga tgccgcttac gaacatcacg 35400
tacgtcataa gctccgtgcg ccctgatcac gctcgcgttt acgaggtttc ggaggtgttc 35460
cttaacagcg cgatgtttgt ctcggccgtt tatccaaatt gttctcattt tacaccgccc 35520
ggtaccgccc tacacattcc gatattgtat aacttttccg ccccgaggat aggatgtccg 35580
ctgtgtgatt cgattgtttt gagctatgat gagaaccagg gactacagac catgatgtac 35640
gtgtcaaacc ccacagttca agccaacctg ttttccccgt attctccttt ctttgacaac 35700
gataattttc atatccatta tttgtggtta atgaacaacg gtaccgttgt ggaaattcgg 35760
ggattgtaca gaagacacgc actcagtgct attgcattag tatttgcttt tattggaaca 35820
atgtctgctc tgtattttct gtttaagctg ttttccatcc tggcctaaac tgttaataaa 35880
gcgtaaaact ttaaaaggtt gtttcccgtt tctttttgga tggttacgcg gggactgaga 35940
ttagcgagaa cctcggatag aggcgtgccg tcacatattg attcaatcac ctctgcgcca 36000
acttgaatgg ccagctctcg atttgggga ggaaacatta aaaaatattt gtcggtcgcg 36060
gttaataatc gctgcggagc ccgcgacaaa actcgaaaat ccttctgaaa gtcgtgggac 36120
tctatggagt ttaaacaaaa caccatgtct tccaggcgct gttttacgcg cgagagcatg 36180
gattcgaagt attctctctg atacgggccc gccgccttac acgtattcgc gtatagcatt 36240
ttaaagtgaa acagaatatt ttccaaaccc aaacgctgca ataggacgtg ctttaaaaaa 36300
aggtgttgga taggctttag ttttaaatta ggtaaaaaca ttatagagtt gttggcgcgc 36360
tgcataattg tcacggtgtt aaacgcggta atcgggctct cgcctatttg tggctggcgc 36420
gaacccatgg caaaaaatat gtctgccagt aaacagtccg actgtacatt tgactcgctg 36480
gtcgctggcg gttgcttgag gatggatcga agttttttat tatccggcgg aagcgcgcta 36540
cataaggtcg gtggcgcatt tggatcgtat tcttgcgtgg tgtttgccag ctcgtttatg 36600
gttatctgtt tgttatttgt aagagttacc ggtggggtct tgacggtcgc cgggacggaa 36660
aatggcctgc agtgcaaaaa aagagagtta ctcgcaatca ctaattttga gtttacggta 36720
attggtgtgt cctcaaagtc agaacttcta attactgggt ctctgaacaa tgatacgcct 36780
ctctcgcctt cgtccgttcc cacgcacaag cagctgttta gcgcgtaggg atggggtgac 36840
accataattc gcagcattga taattccgga cgttttgggg cgtacaatgg catcatgttt 36900
gaatagatgg tgctgaattt tttaaacaac aggggaaggg gtacggccgg tcccggaagt 36960
ccgaccattg cggcaagtgc cgacgagtac acgacatgac agacgtactc gccgttggat 37020
agctgcaccg cccccccaac taagcgagcc tttaaatctg agatttttat cattgatttt 37080
cctttggtag tttcaggaac cgcagcggcc ttggcttcgc ggttatattt ttgctttatg 37140
tctccagtga agtcccaaac gtctaatagt gtgtcgaatt acctgcggtt ttaagtttga 37200
aaacgcgtga ccggtcgtag ctatgaaaga cataaaattg gcgttgcgtc ctaactcgtc 37260
taaccgtatt tttttacttc gcacgtatat ctttggaatg aggttgcatt ttttatatcc 37320
cagaatagcg cattcgtgtt tgcttttggt gtgactaatg atgcgttttg aaaagtcaaa 37380
cgtaccaagg cgggcggcgc ggttcagggc tcccagatcc aaacgcaaaa acgaacaaca 37440
ttccaccaaa cactcgataa aacgcttgtg tagagtttgt atttgcgctt tatttcttgt 37500
ctccagcgtt gtgaatattg tgtttttcac ctggtctatt ctagcatatg gaacgcttaa 37560
tattgggttc aattcatccg aaagtttttt aagtaagggc ctaccgtata ttttagaaat 37620
tgtcgcagtt gccgccttaa accaggaaac gttgcatccg tctttactag aaaccacaga 37680
agggaaatta gtggtccaga aaacgtctcg ccgcccccgt ctaactatgt attggttttc 37740
aagcccgtgg aggtctaacg tagagttcca gttcgcaatc gaaaccggga aaatgtgctt 37800
caccggtaaa aagcacctga ggtttcccat agcgaacata aagtgcagca tgtccccgct 37860
tatgttaacg ttaattgcag ccccgttgca caggttgtcg gagtagaaca cactcacggt 37920
gaatgacggt tctttcacgg agaatttttt cattacaaag ttattagtga gcctggggat 37980
gtcgataatt gactcgaagg taccggcgag agaccagcac gtgattttgg tattaattac 38040
catattggta ttgaaaactg agattttgaa gttgtgtaaa ctcatgtgct taactggatc 38100
gaacatgtct ggaagaggtt cggtctggtg gtcgtgacgc gcttgttcat atcgatctgg 38160
gttttgatcg gcgttaaaaa taaaattgta atctctatct tttggtgcgt attgaattcg 38220
cgaaaacatg gcctctagga cgcgaggcgt ttcgcctggc cccagaaact cggtgttagg 38280
atacagcctg tcgcacgaat ggcacacgtg ttccgccaca gacatgcaca ttgcgacgcg 38340
gaatattggt acctcgggtg gaaacgttag atagtacgtt gactttggcg tcacaacaac 38400
taccttgtac ggtttttgc cgggatcctt tttgttaatt ttcacctgta gatttacgtc 38460
aagcgattga tttactatgt tttggtagtc tatgattaat tcgctcgcca ctttgatgtc 38520
ctgaaatttg ttttgctttg agaggttttt taaaagtacc ggattggatt cgatgttgtg 38580
aataattacc gtttggacgc gatgcagttc gggatgcgaa atccaatccg aagtaaaacc 38640

caaaatgtta aattttaagt aattcagatg cagttttgca cacgggggcg gacaccgagc 38700
tagcgtgagg attttgaatg gctctaaaag tttaaagtaa ttctgataac acgtgccgat 38760
cagatgttcc ggcagtgtgc gcgtaaaggc agcaaagcat tcctcgaatt cccatatttg 38820
aaggtaacga cagaggattc gaaacatgct ttgataattt acgccacata ccattaatgg 38880
cgaacagtac agcgaccgta ggattcgttc gtcgctttga accaaaaccg tgtcggacag 38940
ttcgtaggtg aaaacggtgg gtaaaaacgt atgcgtctcg atatgggaaa atgtacatcc 39000
gatagcatca tcaattatta aacccggtcc cgcgtactcg ttaccggtct tgccaacgat 39060
accaccttgg cggtgtataa aagggtatag tccttctccc aataccagac gccggagcgc 39120
acggcacgcg taaaggtgg tctctgtttg catagccctg tccaaggcct ctagcgagac 39180
gttttcaaga atctcattat tggtggccag gcagcaaaaa attgccagtt gagttttaaa 39240
atcggaattg gccggtgacg gggcgttcgc tgccaccgta gcaggacacg cgggcagaag 39300
aacagggccc tgcaacagca taatggaggc ggctctggag gtacgaccct ttccatatat 39360
ggccacggag gccaatctgc ttcgtcaaat gaaagagtcg gcagccagcg gactttttaa 39420
aagttttcaa ctgctgttgg gaaaagacgc cagggaggga ggagttcaat ttgagggtct 39480
tttgggcgtg tataccaacg taattcagtt tgtgaagttt ctggaaacgt cgctagcggt 39540
tgcgtgcgtt aacaccgagt ttaaggacct caaacgaatg acggatggaa aaattcagtt 39600
taaggtatct gtaccgacca tcgcgtatgg ggacggcagg cggcccacaa aacaaaaaca 39660
atacattatc atgaaggcct gcaataagca tcacattggt gccgagatag agctgtcgac 39720
tgatgacatc gagctgctat tcattgacag agaaacccca ctcgattaca cagaatacgc 39780
cggggccgta aagacgatta ccgcctctct ccagtttggc gtggacgcgc tggagagggg 39840
cctggtagat accgtattga atgttaagct taggtccgcc ccgccgatgt ttattctaaa 39900
aacactatca gacccggtct acaccgaacg gggtctaaag aaggctgtta agtcagacat 39960
ggtgtccatg ttcaaaagct accttatgga taactcgttt ttcctcgaca aatcagacat 40020
cgccgtcaag ggaaagcagt acgtgctgtc ggttctctcc gacatggtgg gggcggtgtg 40080
tcacgaaacg gtttttaagg ggacgaatac gtatctgtct gcatcgggag agccaattgc 40140
cggagtcatg gagaccacgg aaaacgtaat gcgaaaactg ttaaacatgc taggtcaggt 40200
tgacgggggc atgtccggtc cggcgtctta cgccaattac gttgtcaggg gcgaaaatct 40260
cgtaaccgcc gtgacgtacg gtcgcgtcat gcgaacgttt gaccagttca tgaaacgcat 40320
agtggaccgt cccaacgcgc agcccagcgt cgacgatgat cgggacgcgg tggcggacgg 40380
gcaggactcc ctcgccaaaa caccgatcgc ggcggctgtt atccagattg gagataaact 40440
ggtggctttg gaaagtttac agcgaatgta caacgagacc cagtttccct tcccgttaaa 40500
caggcgtatg cactatactt attttttccc cattgggttg cacatgcccc gtcctcagta 40560
ttccacgtcg gccacaatta aggggtcga gcacccagcg gaacaatccg tcgaaacgtg 40620
gattgtaaat aaaaacaacg tactgctgag tttcaattat caaaacgcgc tcaagtccat 40680
ctgtcatccg cgcatgcaca acccgatgcc gtgtgggcaa gctttgggtc aggcgtttcc 40740
agatcccgga cacgttcaca ggtacggaca aagatctgag caccccccga acatgaacct 40800
atacgggctg gtgtacaact attaccaggg caaaaacgtg gcacacgttc cggatgtcgc 40860
cctaaaggcg acgatgacca cagatgagtt gctgcaccca acctcgcacg agacgcttcg 40920
tttggaggtt caccccgatgt ttgatttttt cgttcatcag cagcctggtg cgcaagccgc 40980
gtatagggcg acccacagga ctatggtggg taacattcca caacccctgg cgcccaacga 41040
gtttcaaaac agcagaggcc tgcagtttga cagagcggcg gccgtggctc acgtgctgga 41100
ccagtcaacg atggaaatta tccaagatac ggcgtttgac acgtcgtacc cactactctg 41160
ttatgtcatc gaatgcctca ttcacggaca ggaagacaaa tttttgatta attctccttt 41220
aattgcatta accattgaaa cttactggaa caatgccgga aaactggcgt ttattaacag 41280
cttccctatg ctgcgattta tctgcgttca cctgggcaac ggtagtattt ctaaggacgt 41340
gtacgcccat taccgaaaag tttttggcga actcgtggtt ttgcagcagg cgctctcgaa 41400
aatcgcgggc cacgaggtgg tggggcgcag gcccgcgtcc gagctgatta actgtcttca 41460
ggaccccaat cttttgccgc cctttgctta caatgacgtt tttaccaacc tgctcaggca 41520
gtcctcgcgg caccccatgg tactcatagg cgacgagggg tacgaaacgg aaaatgacag 41580
ggatacgtac atcaacgtca gaggaaaaat ggaggaccta gtcggtgaca tggttaacat 41640
ttacgagacc agaaacaacg cggatcatga cggccgccac gtccttgacg tcggtccctt 41700
taatgaaaac gaacagcaca tggctgtgct ggaaaagctt ttttattacg tggttctgcc 41760
agcctgtacc aacggtcacg tctgcggcat gggcgtcgat tttgacaacg tggccctggc 41820
cctgacgtac aacggcccag tgtttgctga cgtcgtgaac cccgacgatg agattttgga 41880
ccacctggag aacggacgc tccgcgagat gctcgaggct tcggatatac accccaccgt 41940
tgacatgatt cgaactcttt gcacgtcgtt tctcacctgc ccgtttgtta cccaggcctc 42000
ccgtgttgtg actcagcggg accccgcgca actgttgacc actcacgacg acgggagata 42060
cgtgagccag actgtcctcg ttaacgggtt cgcggcgttt gctatcgcag ataggtctcg 42120
tgacgttgcc gagaccatgt tttacccggt gccgttcacc aagctgtaca gcgatcccct 42180
ggtggcggcc acgctccacc cgctggtcgc aaattacgtg acgcgcctgc cggcccagcg 42240
cgtgccggtc gcgtttaacg tccccccggc cctcatggcc gagtacgagg agtggcacaa 42300

gtctccaatg ctggcctacg ctaacacctg cccgatgacg cccacgtcgt tgagcaccct 42360
ggcgagcatg cacatgaagc tgtccgcgcc ggggttcatc tgccacgcaa agcacaagat 42420
tcacccgggc tttgcgatga ccgccgtccg aaccgatgag gtgttggcgg agaacttgct 42480
atttagtgcc agggcctcga cgtccatgtt tttagggcag ccatcggtta tgcgtcggga 42540
agtcagggcg gacgcagtca cgtttgaggt gaatcatgag ttggcatcgc tggacatggc 42600
gctcggttat tcttccacca tcacgcccgc ccacgttgcg gcgattacct cggacatggg 42660
cgttcactgt caggacatgt ttctcatgtt tcccggggac tcgtaccagg acaggaccct 42720
caacgactac gttaaacaaa aagccggatg ccaacgattc ggtggtcctg gccagattcg 42780
tgagcccgtc gcttacgttg cgggggtgcc gcactcggac aacataccgg gtctcagcca 42840
cggacagctg gccacgtgtg agattgtttt gacgcccgtt actgcagacg ttacctattt 42900
tcaaaccccc aacagtcccc ggggacggc atcctgcgtg atctcgtgtg acgcgtacaa 42960
caacgaaagc gcggaacgtt tgctctttga ccactccatc ccggattctg cctacgaata 43020
ccgcactacg gttaacccat gggcgtcgca gcagggctcc ctcggagacg tgctgtacaa 43080
ctcaacctcg cgccaggtcg cagtgccagg gatgtacagt ccgtgtcgcc agtttttcca 43140
caaggacgct attttgcgta acaatcgggg cctgaacaca ctagtcacgg aatacgcggc 43200
ccgcctcacg ggaacgccgg cgaccagcgc gacggacctg cagtacgtgg tggtcaacgg 43260
aacggatgtg tttctagaac aaccgtgcca gtttctacaa gaagcgtttc ccacgctcgc 43320
cgccagtcac aggtctctgc tggacgaata tatgtcgaat aagctcacgc acgcccctgt 43380
gcacatggga cattatatga ttgaggaagt ggcccctatg aaaagactat taaagatcgg 43440
aaacaaggtc gcctattagt ttagctcaga cggtctggag ctaacgagag atggccctcg 43500
ataagagcat cgttgtctcg gtgacgtcta gattattcgc cgacgagata gcaaatcttc 43560
agtcaaagat aggatgcatt ttgcctctca gagacgccca ccgtctgcag aatatacagg 43620
cgctgggtct ggggaacctg tgctctaggg attccgcggt ggattttatt caggcatatc 43680
actatttgga caaatgcact ctcgccgtgt tggaagaggt cggtcccaac agtttacggc 43740
taacgcgcat tgatcccatg gacaattatc aaataaaaaa cgcgtaccaa ccggccttcc 43800
attgggataa ctactcagaa ttggtagtta taccaccggt ctttgggcgc aaagatgcga 43860
ccgtctcact ggagtctaac gggtttgatg tggttttccc tgccgtggtg ccagaaccac 43920
tggctcaaac agtgcttcag aagctgctgc tgtataacat atactacaga gtggcggaga 43980
cgacgcccac cgacgtcaac ctagccgagg tgacgctgta cacgaccaat atcacttaca 44040
tgggtcgcaa ctacgccctg gacgtggacc ccgttgggtc gagctcagct atgcggatgc 44100
tggacgacct gtccatttac ctgtgcgttt tgtccgcgtt aattccgcgc gggtgcgtaa 44160
ggctactgac ctcattggtg cgccacaaca aacacgaatt agtcgagatt ttcgaggggg 44220
tggtgccacc tgaggtacag gccctggatc tcaacaacgt aagcgtggcc gacgacataa 44280
cgcgcatggg tgccctcata acctatctac gaagtctcag ttctatattt aatctgggcc 44340
gcagatttca cgtttacgcg ttctcatcgg acacgaatac cgcttcctgt tggtgtgcat 44400
ataactagaa acgggcctcc ctgtgctttg acatgtcgat ccccaaaatt atgacggtgt 44460
ccagagacaa cgagggtacg gtgtgtgaag tcgcggtgga caacggacga cacagagcga 44520
tgatttatta ccctaagacc accaacttag caaacgagcg cgcggacgtt gttaaggaag 44580
cttttgatac cgaaacccca gtggacattg taaagcaaat tgttaacgag ggcctagcta 44640
tatccaaaaa aaattgcgtc cgtttggcgt tgtatttata tttttatttg cagtacgtgt 44700
gctttgctct gctcctcact tggcagttaa acccgtacat ggacccaccg ggtctggtgt 44760
ttgcggttaa ccccatgggt ccaaaacatg tcacgaaact accgcacccg gctattgttg 44820
cggtaggttg tggggcagac gccatctgta agaactgtag cgtccccgat atcaaaacgg 44880
agcttggaat ggtttaccac aacgggtcta gcgattctgg tcagcgcgca cactatgggc 44940
tggccctgtt aaaggcggcc tggcttgtca tgggaaatgt gtgtccggaa ccagtagtgc 45000
ggcaaggcgc tgcattactt ggtccatgga accggacgga gtggtcggat tttaaatcgg 45060
caatggcggc aaccacgttt tgcggatcca gaggcgttct gtggtcaccg attcatgaaa 45120
aaaacctctg tcgccccacc tggaatgatg taattaacac atcagttttt acaaatgaat 45180
cactctgtcc aaatatacct gtggtgcccg aaagtgtaat agtgcttaat ggtgatgcat 45240
gagaacaata aacgtattcc cacgcacttc atgtacgttg tttttattgg tctggtaagg 45300
tattaacagt aatgggaggt tccgctggtg cctataaagc aaaacggtca tagagtaaca 45360
atattgtgaa tggggaaaat agtctgtaaa tagtttatcc gaacagtatg actgcacaca 45420
cgaatggggt tttaaccacg acgggctttt caacaagtca gccggaatcg gttcaagttt 45480
ctccatttta tcgcgtaatt acaaaacctc ccgttatggg cttgtttttt tgcgtggcta 45540
tgtgcgttat cgcgttggta tggtacgtga tgcggagggt gtgttgtaag gggcgcgttg 45600
ttgccgattc gtgtcgcgac ccgcgtcaac ccgcgtatga gatgttgaat gttaggttgc 45660
gtccccacgg aaccaatcca tagagaactt tcacgtacat agccaaagcg cacgacgtgc 45720
tctctagctc gtttattcta gcgatttaaa aacaaattta tcgttggttg ccatgtagtg 45780
ggccatgacg gcggcgacaa gcacgtcgtc cgacatggtt ttttgtttcg cgcagtacgt 45840
gtgacctccg tcttttagtg gaatgcaacg tattgccttg atttgatcta ttagatacgc 45900
tacgggatca aacgatagct taatagtgtg agacaccaca gtttggctag cgctaaacgt 45960

tcccgagttc aatgcgtaaa taaaagattc aaacgccttg gctttctctg gccccaacat 46020
gtaaatgggc gaacgtataa gagtgttctt gtccacgtgg tgtaaaaaac taagagggac 46080
cgagcaaatc tcgtttaaca cggttgctat ggccactccg gcatcttgac tgctgttgcc 46140
ttccacggcg acgttaacgt gcgtgatctg cggtgaaggg gtaacgatcg ctcgaattaa 46200
cgcggcggca catgacgcta tctggtacgc ggcggtgccg gttagatctc ttagaaaaaa 46260
atgttccacg cctaataaaa tgcatttaat caccttgtga tttaccgcaa tcaccgcacc 46320
gatgccggtg ccggacgcgt ctgtgttgtt ggtgtacgcc gggtctatgt aaacgtggag 46380
acacggcttc atcgcgccct gtatatcttg tgacgtgctc tttacgcggc acagatccaa 46440
ttgtgacagc gaagagtcgc taacgatttt atgcatgctt tgagcggacg tggcagcgtc 46500
ccccattagc tccgtggaga atgcgccttc caaaaaaaga ttagtcgtgc tccgtaccgt 46560
ctcatcaatc gtaatatagg caggaatatg caggcggtag cacgggcacg ccacgacagt 46620
gtcttgcaag ttaaaatcat ctttatgatc cggacatacg tagttcacca cgttaagcat 46680
cttttcgtgg gcgtccttca aattaagtaa aaaactagtt gatttgtcag atgagttact 46740
tgaggatata aaaatcagtt ttgcgtcctt ttgaagcata aaacccagaa tcgcggggag 46800
ggcatctttt ttaatgaagt tcgcctcgtc gatatatagg aggtggaacg tctgtccccg 46860
gatgctctaa aggcaaagaa aaacacaaaa cgagttacgc gggtatggag aacgatacgc 46920
ctaaggacaa aatctcggaa gctgactttc aacagtgtca ggcgttcttt caccgtccca 46980
ttagagatct aatttcatct ggagctgacg ctttaaacca ctttagccta tctgaatcag 47040
acggacataa attggaacgg attgttcttc tgcttgacct ggtggggaca gaatgtctct 47100
cttataccac gatcgctgca aagaatgtca aatgacgcgc gtcaacagcc caatatgtcg 47160
atttcataac gtctctaact tataccagtg tttggattgt aagcgctatc acgtatgcga 47220
cgggggacgc aactgcgtga tcgtgtacac tcgcgaaaat ctagtgtgtg atttaacggg 47280
aaactgcgtt ttggataatg tgcaggacgt atgttcgtac ggtcctccag aacgccgcgt 47340
acccgacgcc ttcatcgatc cgctcgtgtc acacggcacg agggaatgtc ttaaaagcga 47400
tatactgagg tactttgaga cggtcggtgt gaaatctgag gcatattcta ccgttgtcaa 47460
gaatggacaa ttgaatggca tcataggtag attaatagac gctacgttta acgagtgcct 47520
tccggtaatg agcgacggcg aaggtggcag agacctcgcg gcgagcattt acatccacat 47580
aattatctcc atatactcca ctaaaacggt atatgataat cttctattta aatgtacgag 47640
aaataaaaaa tacgaccaca ttgtaaaaac tatcagagcg caatggatgc gcatggtctc 47700
aaccggcgat ccgtcgcggg tcagtgcgac gggttgtttc acgtgatact tccgcgaggg 47760
tttatcctcg cgaacaatat tacgtgcggt gaacggcagc ggttttttgc acacacttgg 47820
ttcgctgcat ctggacgcac gtctaagact ttatacgtgt ggggacgggt atttcaaaac 47880
accgacccgg gccgcgggga cggtccgtcc gggccgtggt ccggactggc gattagtctg 47940
cctctgttta ccacaaatgg aaaatttcat ccgtttgatg tagttatact caaggccgat 48000
acgcctgact ctggaagctc gtggaccgtg aagttcttgt atatgtcatt aattgcggct 48060
tacagaaacg caatgcgagg tttaaaagat aaagtttcgc aatgtaccga tgccgccgtt 48120
gacggtgagg ttcatcctct aaccgtctta aaagaagcgt tggtatcacc ggacactgct 48180
acgcgacccg tgtccgcgtg caaccctcta cagatgttga ccggactctt acagtctagg 48240
gtacgggacg actacgtgac acaccaccgt gcgctcgaac gcccaggtaa tgtgagggga 48300
caagtaatcg ccccgacgcg caccgagatg ccaaacggat cgccaagtcg tgtaaggctt 48360
ggattccgcc ctcccaaaca agccaactat ccaaagacgt gggcgcaggc gcgtcacgtt 48420
ttctcgtctc gcgcatatta cgtgtgcgta tatgataacg aagaactaga taccaagtgg 48480
cagcggcaag atccgcggcc gttgccgcta gattggtccg atccggtcgc gtacctgtta 48540
gagggcgatt tgtttttagg agccaagcag aatgcgtttg tcgattctat agaaaagacg 48600
tgcaggtgtc agaactatac cattaagcaa ttttttccgg ttttgataaa tagggacaac 48660
gaaacagtcg acttaattaa ggagcatttt atagaggcgt gcttcgtgat tagaaaccag 48720
gtgtcagaga ggagcgcttg ggtaaaggcg gcgctgtttc gcaacgatag taacacgtat 48780
tggaaggatg ttttgggatt atgggagcat gggcctcata agctgggtac ggctataaaa 48840
ctaccaacat cggaaccttg caatgccgac gtaaactgga gctggctcct gtgtgacgag 48900
gatataactc ggtcaattag cgggcagtct actgtttgct tagttgtctc gcctaccctg 48960
accgcctggc tggtgctccc ggggggcttt gttattaaag gccgctacga cctatcaagc 49020
gaggatttaa tgtttgtggc ttcgagatat ggccacccag cgtcgtcaca ttcttaaatc 49080
gtttttaaac aaggaatgca tatggttgcg acacccgggt acgtccgcgt ttgttcgggt 49140
atacaccgcg accactgcgc attctgccgt ttttgacccg ccggtaacta gcgaaaatgc 49200
gatgtcactt aactttttaa atgttatgat cgtaatcatg aaaccaaaag aatttggccc 49260
gtgcgttacc gtgtacatga atggagatat tctagatttt tgtgccacgg aatctgtcgc 49320
cataagggac gtgcctggta gggcggacct gtgtttaatt cgttttggta ccctttctaa 49380
tgcgccgagg agcgttccga tacccgggcc gttgaaccca catccgcgag aaaccgtgcc 49440
cgggctaaca aaacaggaaa ttatatacac ttcgcaaaca gtgccaagag gacagatacc 49500
agatgccata aaggggaaag agttccacca aataaatccg tttttgtggt ttgacggagg 49560
ggcgttttgg caactgttcc tctctgtgga ttttatgctg ctctgtcccg cactcgacac 49620

agttccgtcc ctggccagaa tcgttgggct tcttacacag tgcgataaga gcacgtgtaa 49680
aatttgtacg ggggcccacg tacacgttaa cccgtatcgc ggatacacgc cacctgactc 49740
gcaagggacc tcaccctcgt gcccctgcct tatctcgtgc ggggccaggc gcgcggcgga 49800
tgtcctggtt accggacacg ttaatctttt gggcctgctc tttgacccca aagcctcccc 49860
caaagtgacc aagctgcgtt taaaaagaaa cccacgcccg gtaccgatag aggacgccat 49920
gtcgggcgtc acggccgaag ggaccgaggt gcaacccact tcgctaccgt gggccctcat 49980
tcgcctgccg gatttagcca gtcgcgtgat gctatacggc tgccagaact taaaaagcat 50040
ctgcttacgt tcttattgaa gcacgtcgca cacatcagcg aactcggccg ccgtccgggg 50100
ctcgcgtaca cgatggttcc gtttccttc tttatgttca ggttcttcgc cggaaaccac 50160
ctggatagcg tggcgatgac ctcggagaaa acggcattgg cgacgtgttt ttggtgggcc 50220
acgtaaccca cgtgaacgtt ttcaaccgac gagagaagaa cgctgataat ggcgaccacg 50280
atccatgttt tcccgtgacg acgcgggata acaaaaacgc tggctttctg cttaaacgtc 50340
tgtaacacgt cgtccgtggt ttcaaataaa ccaaagtgtt gcttaaacgt ggcaaataac 50400
tggtgtgtct tttcgggtgc tttgatagac gcgataaaat aaaacgtgtg catgattaat 50460
tgttgttgaa agggctcaag gcaacaaacc ccgggcacgt aactcccatt taaaaatgac 50520
gacaggttgg ctaaaaactg ccgcagttct gcataaacgg ggcactctag gaaagcttcg 50580
tgggtccgtt gggccgattg gtattccatg tactggtcct tggcgtttgc catttggacc 50640
cgacaacatg tttgtttaag gtagtttgtg agttcctcgg ttaaccgggg caacgtagcg 50700
tggctagaaa ttcggggctt tcccaacgta ggggccggca tgggttcttg aactctgcgg 50760
tacgtttgaa gataatcgtc taaaacagaa ctatatgcat taaccgcgtg aataactccc 50820
aagcatgggt gggccattcg ttcggtcttg tcgttgcccg atataataac gggcggttgc 50880
cgaaaccagt tttcgcaacc accgtccgtg accacgcgca agttattttg aagacgttcg 50940
cgatagctgg ttaacaacat gttcccttca agtttttga ataacgggca ccccgaaacg 51000
gaacgccgct tcgttaaagg cgttcagtta gctttagacc tgtgtgacaa cactccggga 51060
cagtttaaac tagttgaaac acctcttaat agttttcttc tggtatccaa cgttctgccg 51120
gaatcgcgcc cggttagaga ctgtccgcag ccggaagggt ttgactttga acacattcac 51180
ctcccaaaac taacacgcat gcagcgtgtc ctggggcgat actgcgacca tgttaacaac 51240
gacgacacgt gcgttaacgt aaaggcaagt tcctcgaatt cacagggtgc cttgttttat 51300
ctgccgtatg gacaggacga atggaattgg gcgctcacgt taaggaaaga caagttggtt 51360
aaaatggctg tagagggctt gtcaaatccc acgacctgga aaggtttaga gcccgtggat 51420
cctttaccgc tcatatggct tctgttttac ggttcccggt cgttctgtcg ggaaccagag 51480
tgcctatatg aacgcaattt tggtatgaag ggacccatac tcttaccgcc acatatgtat 51540
gcccccccaaa aggacgtaat gacttttgtc catcatgtaa ttaagtacgt taaattttta 51600
tacgtgaacg ccggtggggg tcttgaaact gaaccgtccc cgccgttcga ggcctcgcgg 51660
ttgcgcgcag ccatcgctcg tctcggggac gtggaagcgg atgacgcata cctgtccgca 51720
aagtgcatgt tgtgtcacct gtacaagcaa aacgatacga tttcgattca tgaaacacac 51780
gtgggcggag tcatcgcatt aggcggagac ggtgcgagat atataacgtc tagtgttcgg 51840
gctcaacggt gcacgagtcg gggagatttt gttttaatcc ccctgtacaa cattgaaggg 51900
ctcgtaagca tgataaggga acatggcctc ggcagcagct aaaaaaatgt taattaagtc 51960
ggagctcgag tcggaaatca acaaaaaact gtccatctcc gtatttgaca ggtttggggc 52020
cgacagtgcc gtgtttaacg cgcagtataa gggaaccagg gaatcgctgc ggtcgtacaa 52080
cagcctaaaa aagaaggacg atctggcgac cgttgtcgga acgctagaaa cgtcgctgcg 52140
tgaaaaacaa agcgaattgg gattactaaa ggggtttaac aggaaaaaaa ttgaagagtt 52200
tgacgctgtg gcggacgcgg ttcgcgacct caaggacgag ctgtacggag aactggagat 52260
tctaggtacg cttgacaatg aatctgttcc cgtggaagaa gagtccccca aggacgacat 52320
tattaggtgg aaattggagc gtctgcccag agtgtgcccc aaaagccctt gatcccattc 52380
ccaaggttca gactgacgtc gacagaacag catcgtccca tataaccgtc attaaaacac 52440
gtaagacgat cgcccaactg aagataccta acaactgggg ccagtgtagt caccaggcga 52500
cggactggac cgccgtgctc ggacgcggct cgtatggtgt ggtgaggtcc atgtctctcg 52560
gccgctgcgt taagcatttt ggcagccggc gtgagttttt ttacgagtgc atttttaacg 52620
atatagtacg cgcctgccgg gagaaacatc ccctgaaccg cgggggtgac cgtatactat 52680
gtttcctaga gccgtgcgta ccatgtcgcg ccctgatatt cccgcagtta acagggaatc 52740
tgctaaacgc ggatcttaaa cacgtgaacc ctgaacggct ggccgttgaa ttctctgagc 52800
tgagggaagg cgttagtttt ctaaataata tatgtggcat cgttcactgt gacatcagtc 52860
cagaaaatat actgataaag gggggaactga caactgcgta cgggagactt atgatcggag 52920
atctagggtc cgcctcttta cacacgggaa ccccttggac cggagtgatg gtgacctcca 52980
aactcgggtt cgtgcagcac acgtaccatt ttaaggcacc ggccagattt atctgtaagc 53040
acatttaccg gccgtcgtgt ctcctctacc ggtgtctgct gtcgtgcgcc gggggcccgc 53100
aggcgcatat gctaaatcag ccgttccaaa tcactccaca actcggtctc acaattgaca 53160
tatcgtccct gggttatagt ttgctagcat gcctagagaa atatcttcag ccagctgacc 53220
catttcccca gcagggagcg ttggcggacg cttcctccga atccgcccac ccattgttct 53280

atttgcgttg catggtgcca agagtagtca tcgccgagat tttttctgtt gcctgggacg 53340
ttccactcga tttaggcatt gactcatctg gccacgcgcc agctattccc ctgagagaag 53400
cgtacaggcg gttttttgcc aaccagtgta gtttatatag ggcgcaatac aaagaggatg 53460
cgttagaaaa cgcatcctcg cggctgtgta actcaaaact taaactagtt ctccagaagc 53520
tgttggtcag ggactacttt agtcattgcg gaaactgcgg agatcatgga tttttttctca 53580
gatgagccga tggttcagga gatggcgctt ctcgacatcg atgagcagca gcggctcctc 53640
tcgaaaatga gcctggccaa ctttttaaaa catgagcgag taagggcgtt ttttagcgat 53700
aacaaaaagg aaataagcat gccggctata cggttcgtgt ataattttta tctattcgcc 53760
aaggtgggag attttatcgg caacaccgac gtgtacgatt tttacgtcac ttgcgtgttc 53820
agggggaggc gcctgacgcg cctgtcggaa gtgtacgacg cgtgcctaaa catgcacccg 53880
cacgatcgac accacgtgtg tgcattgata gaacaggtca cgcgcggcca aaacatcaat 53940
cctctgtggg acgctctgag ggacggcata atttcgtcgt caaaatttca ctgggccata 54000
aaacaacaga attcgtccaa aaaaattttt aacccgtggc ctatagtcaa caatcacttt 54060
gtagcgggcc cgctcgcgtt tggactgcgt tgcgaggaag tggttaaaaa gatactggcg 54120
acgttgctgc atccaggcga ggcgcactgt gaaaactacg gattcatgca gagtcctctc 54180
aacggggttt ttggcgtctc cttggatttt ggaattaacg tcaggtctga cccaaaagac 54240
ggtttggagt ttcacccaga ctgcaaaatc tatgaaataa aatgccggtt taagtacact 54300
ttttccaaga tggagtgtga cccgatttac gctgcgtatg ctaaacttta tcagaagccc 54360
agcatgcaga cgcttaaggg gtttttgtac tccatatcta aaccggcgat cgagtttgtc 54420
ggagaggaca ggctccccag tgaatcggac tatcttgtgg catatgacaa agaatgggag 54480
gtgtgtccgc ggaaaaagag acgcttaact gcagtacacc atctagttaa aaagtgcatg 54540
attcacaact ctacggcgcc ttctgatgtg tatatattgt cagatccgca ggaaaccgga 54600
ggccaaatta atattaaagc tcatctgagc gccaacctat ttataaacgt caggcatccg 54660
tattattatc aagtgttgct ccagtctctc gttgtacagg agtacatcag tctctccaag 54720
ggaactaaaa atttgggaac ccagaaaaac tttatagcga ctggcttttt tagaaaacgg 54780
cagtttcaag acccgagctg ctgcacgatc ggtgaatttg ccccgttgga tccacacgta 54840
gagataccga ccccttttaat cgtgacaccg gtgtattttc ccagcgtggc caaacaccaa 54900
ctggtgaagc aggcgaccga attctgggcg gctagtgctc gtgaggcatt tcccgagcta 54960
ccatgggatt tatcctctct gtgtgcaaac gccccaccaa caccgtagat gtgaaggggg 55020
agcccataga tgtatccaaa gaattcgatc ctattatagg agaagaaagc attgtcttgt 55080
taacggcaga tgggactgcc cccgcggcgc tgtacaaacc caaaaccaag ccatccaaac 55140
ataaaaacaa taaattgtca gatttgtttt aagcattcgt atctttattg gagtggggtg 55200
gatggtgtgg ggttgggaag ggaatgggat tgaggggagg atgaaatgct agaatcatat 55260
gtattttgga tatgcatcct cgtaatcgct cacgtcttct tctgtttcat ataacgcgtt 55320
cagttccggg gagtacttaa cttttgtcgt cttctcttta gggccttgag gaaggggtgc 55380
gtattcgggc gactttgcgc accatttcca atcgtttctg ataattctta agatgatggc 55440
taggatgcaa gtgatcggaa tcatagagat gttcacggct atgactttgt gcaggctggc 55500
agatatgaat accttgtcgt atcgaacgac gggaagaagc agaatcaggt aactgacaaa 55560
aacgccaacg tagaatccga tctggtgctt taggtacttt actagaaata gttctgtgtt 55620
gatataccaa gctattgtaa ggacaaaaaa caagtttatt gaaccaaaga caatatcgga 55680
aacgagcata tagaaactgt ttccaatagc catcatcgtt ccgagggaga atactaacat 55740
ctccatggct attaaggaca gataaagatt ggcgccaatc ggttttccgt aacgcaacag 55800
ggtgtctagc aagctattct caggaatctg ctgctctaag acgcgtagag accaggtggc 55860
cgaattgcac gatatcatag cctggacgtg cgtaaacgat aggcaaaaat gtatacagta 55920
tacaaacgag gctaaaagta tatgtttgta cgacagaacg tgaataaata actggatggt 55980
ccacagcctc aaaatagaca taaaaactac actgtgggag cctatgagca cgatccatgt 56040
ctggaggctg gtcattgttg tcgcgtgcac gcgcttggcc ttgaatatgg ccaggatggc 56100
ccacgcatag taacacagca gatatccgtc caccaaaaac gcccacgaga tgtagacaaa 56160
catttctggt gcttccaaaa acaaagtcgg cgtcagatgt ttagcagaac ttctgaccgt 56220
taggtttaac gcgctgtagt tgaccaacgt gttaaagtag cacggaaatc caagtcccgg 56280
gaatgtggcg gtcagtggca ctaccgctga cattataaac ataagtccaa gaataaccag 56340
tagttttacc catgacgata agataaacga gtcgctccgt gaaattttca tagttcaaag 56400
aatctgctcg cgggtgcgtt cggcaaacta gcagtaaaat cagcaagtgg cttacatgtt 56460
acggatttaa ataggctccg cttttaaata tcacattctg tgcaaaggag ttgagccacg 56520
gcaccatgaa cgcccgggag gtggcactca cgggacatgt tttgcacata tcgctgcata 56580
gcacgcacga gcgcgagaaa ttaataatct ggcaggttca tttacttgta tgtcaacaat 56640
gcggaattca gggagatgcc gcatatctat ttgtcaccga aacattaagc aatactgact 56700
ggggaaatat accggcgata aaccgtcacg caccgtccat aaatgagcac ggccgtaatt 56760
atatgcagtg ggaactccgt actcgtttac ggaatcccat cattcaattg ttaagtcgcc 56820
agcccggtgc ggttaacgta agggtcagcg agccgaatat ggtaatagtg ggctgcgaac 56880
gagcgttgga tcactcgtgt tcggtgcgcg tgactggagc ctatcttcat tgcgatacca 56940

ctatggactt tagtttggat tctgttgtgt ccccaacccg cgaattttgg ttctcagaga 57000
tgttttctca ctgtttagtt tccaacattg aagtttacct taaaacaacg ggcgggttat 57060
actatagggc atcgagtgcc acgcaatgcc gaaaaagggc gaaagatggc gcattgggta 57120
ttcttgatat ctttaattgc gaatctcgtg aaatacaagt tgccgggcag aagtacacct 57180
tgagtatcgc caccgcaaca tttcacgttc tctgggtgga cgaggcgtgt atgtggaacg 57240
gggccctggc cgaatttttt agggcgctgc acaataagtt gttcggcgac cgggaaggcg 57300
tagcgccaac gttaacgtac gtgtgtccgg gggccactcc ggagggaacc cccttccccc 57360
cctacttttc cgcgtttcca cacctcccgc tcgtgtttgg aagaccgcga aggctcgacg 57420
taaccgcggt ccaagaactc ccaaaagcac aaattgcggt acactggccc ccgtttaaag 57480
attcaatctt aggggatcag cttctcatac ctggcatttc acctaaaaag ccaggtaccg 57540
tacccgttcg ttggccgctt tgggtggagg atgttaactt gagtctctgc gagacgacag 57600
aaagcgttgc ccgcatagtc gacccacatt ctatagtaat cataaaattt tcatcactgt 57660
tgtgccagca cctaaaatgc caccgtgcgt ttgttaaaaa tgagttagaa tacatagcaa 57720
ccatctgttc cagcgacctt cgcctcttca tccaagagga atacaaccgg ttacttgcca 57780
ccatttttac gtgggccgcg gcgagcgggt atacctgggc ggccattgat aaaacaacag 57840
tattcatcaa ggctccccag ctcagcgcag ctgtaagtgg tttctgccca tcactaaata 57900
gttgccgtag gaaacaatgt tacgaaggtt aaaaataaca gttcatttcc tttcacagga 57960
acagcaaaag gtcgtgaccc gtcttgaggc gcatttggga cttcccgtac aggaaacttc 58020
ccacccgcct gactggctca agtgtgaggt ctgctccgcg tccgtgtttt taaaaatacc 58080
agccggggtt ttgtatgccg gactcgcaag agaccccacc agggaagcaa aacgggactc 58140
gtggctggac tgtctagtag aaggcgcgac gttgttgctt aacaactcag tgttaccgat 58200
tggggcgctg gcgggtatct tacccaccct ttttgccaac aggcggtgtg ttaatttttg 58260
gctgctgcca cgcgcgtggg taaaatcggc gcccatatgc cctcccctac cgattgactg 58320
tgttacgcct ccacagtttg tcgtgacaaa gcgtggacca atctgctggt acaaggaatg 58380
gccgttaccg gttgacgttg attttatgta ctacctacag gaggcactat gtgtttttag 58440
tgttgtgtcc aacggggagg gtacggagag tcacgcggac aatatacgac aattagagaa 58500
gtttgaaaag gtactatgtt tattttaaaa cagagtctgt ctggggatgc tgtgtaggct 58560
ggttatttga aaatatttta ataccgatgc gttatgaagg tgtctggtaa aaattgtgcc 58620
gcaataaaat atgttaaaag tctccgcgtg gccctcttct tctccaacgt tctctgttag 58680
acgaaaggat tccaaataaa tgggtttcag taagcctttc atgtcgtctg cgttttcgtc 58740
caacagctta tggagtttga accgcactag ggcggaacg cgcaaaatgc cttggggtgg 58800
gtttaactcg tggtatgccg ccaaggtcga cgctatcact tcgatcccat acaatacaaa 58860
cgcggaatca ttacacgtat acctctcgtg taggtacctg gcccggtcaa gcgcggacag 58920
gataaatgat aggggggct tatcgtagtt ttccagcata gaaagaatgc agcagagttc 58980
tgttacgctc gcgctcgccg gcttaaacag ctgcacgcgg tcagaaaaaa cgtgcggata 59040
tatgtccagt gcgtcgatca tgtcctcgtg ttcgaccacg cacgcaaaga caaacccagt 59100
aaacgttgcc gactgaagcc tggtgcgcac gtgcacgcct aaatcaggaa cgtttagtcg 59160
tcctcgaaca cacagcttag gcgtattgga agccatgcaa atgctatttt ttttatgtac 59220
ttccaacatc agtcgcggaa aaactggata attatagggc cctctggacg agcattctgt 59280
ttcaggcatt accgcctcag atcggtgctg ctccccatt agacgcttca gtatttggtc 59340
catcggttcc aggctggcta atttgttcgc aatatttgtg gcctatgtca tttatataca 59400
cagacagcct gctgtccaca tacagctggt ctgcaatatc gtgtaaactt aaggaaatga 59460
gagagtgatt tatgaggtcc atgtgagtca gtttggcgat tatgtacgta aacgggccta 59520
agagaatgga aaccgtgtcc tgcgagtagg cgatgctcgt tcgctgcccc tgattgtcga 59580
ccacgggtgt gaggttgtac gtttgaaaca tttccccttc ccacaggcta gttagctctt 59640
tgagcatctc cataaacgga gggacgtatt gtgaaaaaaa gctgtttact acggagctgc 59700
ctttagggat gtgacattcg gataggttaa tggttgggtc gcgcagaccg gataaagcct 59760
gcaacgtatt ctgagtgagt atattaatat cgtgtctggg cctgcctccc tctggtaaac 59820
tctgaagcaa ctgcgtttca atgcattgaa tcttctttac gtagtcgtct cgttctttct 59880
caaggttaga aattgtatcg aactgctcgt tgatctggtt cgtcaggcat ttgaacacgg 59940
tatttgaaac ctgtttccgg agaccctggg tcgtactttc cgtaccggtc ccaaagagct 60000
tgtttttgtc cacaatgttt tgagagacgt ctgtgacgaa atcctcaaca acgtctgtta 60060
tgccgctcac tgttttgttc tccgcgagct ttatcagcaa gtttaggagc tccctcttgg 60120
ggtcggtttt ttgagacgat tgctctaagc gagccaatat gtctttatac atgtcgttgc 60180
acgtgctacc gaccagcgct tttagaggcg cggtatttag taactgacat aactttgcgt 60240
gctccgcggt cctgtgacat gtcataatcg cggtaaataa ccggtacaga gggctatcaa 60300
aggcgacgcg tccattgtaa attataggcc gtgaaacaac ttcaactgtt atatcttttt 60360
gtctgtaggt gccgagataa gcagccttat ctccgttaaa ctccacagac acctcggcgt 60420
agtctggaat ataaatagac gtcacatcgg tcattatcga ggtaatctcc tgtgccagtt 60480
tgtcgacaat ccgcgatatc gtaacgtctc cggaagcgta cgttttcctc atctgttgac 60540
agtgtgcttg tatcctggcg atcgtgtcgc cgtcgggggt gcgcctgaca ataggcacca 60600

acccaagggc gaccacccag tcgacgtatc gttcgtacga taaaacagtc ctggtgtgta 60660
gtacctgatt aatagtgttc agtaacagtt gatcgactgt tagacgaacg gtacgtgccc 60720
acgattcaaa cgtagaccgt ttgaattccg gcgtcaggtg ctcgacaccg cgccaacggg 60780
tcttcagggt ggtctcgaag ttggtccagt cctttaagac gttgtcgtat gtgcagctcc 60840
ccaatgccgt cttgtacagg tgtatgaaca gctgtcttcc aaagatccct ggatttcgaa 60900
gactggagtg tagcgtttga ccgcgaacgt acgcgtactt gccctgtagg atctcgaaaa 60960
gcgagatgga gagttccgtc ggatggacca aacacgtcga accaaatccg ggttcatct 61020
tgaacatgac gtccgatgcc aaagtcaggg gtgtcgtgga tcacgtcagt cgcctgtcaa 61080
atataactac cagcccaccg gaaatgggtt ggtacgacct ggccttcgat ccggctgaag 61140
actccgggcc gttcttgccg tttaccgttt atctaattac gggaactgct ggtgctggga 61200
aaagtaccag catatcggcc ctgtaccaaa atttaaactg cctgatcacg ggcgcgacca 61260
ccatagccgc acagaaccta tcgcgtcgcc taaagacgtt ctgtcccacg atcttcagcg 61320
cttttggctt taagagccga cacatcaata tagccgtcag aaaagctcat cagaccggag 61380
ccgtatccat agagcaaatt cagcaacagg agctatcgaa gtattggccg gttatagtgg 61440
acattatgaa agaggttatg gcgaaaaaac ccaatggcat gtacgggact atatccaacg 61500
cgaattttga aaccctctcg agaatgaccg gaccgtgttt atggacttcc aatattattg 61560
taatcgacga ggccggaacc ctgtcctctt acatactcac caccgtcgtg ttcttttact 61620
ggttcctaaa cagctggcta aatacccctc tttaccgcca gggggcggtt ccgtgcatag 61680
tatgcgtcgg ttcgccaacg cagacaaacg cgttccagtc aacttacaac cacgggacgc 61740
aaaagacgga gatatcgtcg tgcgagaaca tcctaacatt catgataggg aagaaggtcg 61800
tatctgagta cgtacacttg gagagaaact gggcgctgtt tataaacaac aagcgctgca 61860
ccgatctgca gttcggacac cttctaaaaa ttttagagta taatcttccc attcctgacg 61920
aagtcatgag ttacgtagac agatttgtcg ttcctaaaag taagattatg gatcctttag 61980
aatacattgg ctggacccga ctctttttgt cacatagcga ggtaaaggcg tatctaacaa 62040
acctacacac atgtctaacg ctaggggcg ataccaggga cacgaagctc tttacctgtc 62100
ccgtggtgtg cgaggtgttt gtgaagccgt ttgaggaata caagcgggcc gtcaacctca 62160
ccaacctcac cgtgaccgag tgggtgacaa aaaacctctt taagttaagt aattattcgc 62220
agtttgtgga ccaggacatg tccatagtcg ccacggaatc cactgaacgt tcgacccagg 62280
ttacctttat caccaagttt gttaaaaaca gccacgtgtc tttaaatgga aaaacaaaaa 62340
aatgcatatg tgggtttcag ggtacgtact ttgagttcaa aagaatccta gacagcgaac 62400
tcttcgtgga gacgcattcg caagaccgtc cggaatacgt atacggtttc ttaaacacgc 62460
tactgtacaa cgccatgtac tcgtttcacg cgtacggcgt gactaggtca cacgagaaat 62520
atctgcaaga cctaaaattt gcacccctcc cggccgctct ggcaaccggg cgtgtagacc 62580
ttcaaacggt tcgtgaagag ttaaacctgg aagacgacat cttctaccac gtgtgcagtc 62640
ccccgccacc cgcgggtatc acctccctcc aggttttggt cgacacgtac tgcgccctaa 62700
aggacgtgtt cgcctccaga ataaaggtgg cgtgtcgctg gtttggcggg gagtttgaga 62760
aggaaacgtt ttccgcgttt acggttaaca tggtcgtaag ggacggagtt gactttgtct 62820
ccccttcaga acgtctcaac gggctgttgg cgtttgcatc gaccgttgaa tcgtataaaa 62880
ttaaggggta cacgttttta ccggtagcgt tcggtcgctg tcagggtttg ccactcagtg 62940
atgacctcag gaagaagatg ccctccctgg tcgtgcagga ctctagcggt tttatcgcgt 63000
gcctagagaa taacataacc aaattgaccg aaaccatgga ggacgggagc gttttccaag 63060
tgtgctgtgc gggggactat ggggtcagct caaatttagc catgaccatc gtaaaggcac 63120
agggaatgtc gttggagcga gtagccgtag tatttgggtc ccacaagaac gtccagacaa 63180
gccacgtgta tgtagcaata tcaagggctg ttaactcaaa ttatttggtc atggacagca 63240
accccccttaa aaccctcctc agagaaccag tcgataacac ctccgccaag catatagtcc 63300
gcgccctcca caacccaaac acaaccctca tctactaaat aaataaaaac aagcaaatgg 63360
aaaacattgt gtttttattc agtccaacca cggccacgga tagttgtcat ttccacacac 63420
cggggggtg gccgccatag tttgacgacc agaaacgccc gggccctctg ggttagaagt 63480
agccccgccc cccgtggtcg aggtcatacg tctcttaggg ccctgggttg taccaacata 63540
aatcacgtcg tcatcaccgt tgcagtccct gcgtgacgcg cgcctcttgc aggcggcacg 63600
aacaccgccg cttgtaacgt gatctgcatc gtcgtcgctg tctgaagatg aacacaggtc 63660
aatcacatca gtggcgcccg tgccctcatc ggcgtctgat tcgtcccatg tctcaacaga 63720
gcctcggtcg tccacttcgt cgtcatctga aataataatg gtctcttgtc ctgcagttga 63780
tccagcaaca acaggttctt cttcatcgga agaccatccg ccctctgccg gtgatctaga 63840
tatatcacct gaccaggcgg gcgaggcccg tggcgaaaag ccgtcagctt ctccgccact 63900
ttcatcagtg tcctcgtagt catctataaa aaaatcgtct tcgctgtctg attcagatga 63960
ctggggtgcg cggtgttggc cagtaaccac cgcaacgggc cgcggaccac caaacgcggt 64020
cagtggcgcg ggcgcgctcg gtggtgtggt ggcacctgtc gaagtaatga cggtaggtag 64080
ccggccgctc tggttaaagg actcaccgtc ccgtgcgtgc ggatctattc taggtgggga 64140
ggagaccggc ggttcgtctt cttcatccac atcactgtag tcgctgttct ccgaatgaga 64200
gtcggtttcg tctggcgttt ttggagtagt gggaggggaa tcttggtcaa aaacatcctc 64260

aggtgcagtc atgtcaatca ctggggggcc tccgtacccg ggcggtaccg gcaaccccgt 64320
acgggacccg gcgaacaaaa aacgtcttgt tctctttttc ctaggtgccc cgggaatcgg 64380
cagcatcctg ggagtggccg gcggggtcct tgacgggtcg gataagaaca tagccatggc 64440
cgaaacgttc acctgtaaaa cgcactgcct atccccgatc agttaatatt ccagtcaacg 64500
gcacccctcc gattctgaac tagatagtca ttcgcaagtt taaaatggtt gcagcctaga 64560
aactgcggcc agggagaacc ggttgccgcg tgggcggtag ccaggggaga cggatggcgc 64620
gccttgagga cgagatgtct ctgtccgttt atcagcgtcg ccttgtctat agcctttctt 64680
ccccacagca taaaaacgca cctctgaagc ttggcagaaa ggcagcttat tatgtagctg 64740
gtgaaccaat cccagccaag gttggcgtgt gaccccgcct tccctctttc caccgtcaga 64800
atggtgttta aaagcagaac tccccgtttt gcccaacagt ctaagcaccc gtgagaagga 64860
gcggtgaacc caggtacagt attggctatc tctttaaaaa tatttttgag gcttggaggt 64920
atactgtagt ccggagcaac gctaaaggct agtccggtgg cttgaccgcg atggtaagga 64980
tcctggccaa ggatcacgac tttaatatcc tccggttcgc aacaatacga ccaccacata 65040
attctatcta tcggtggata aatcacggtt acgttgctca tatccataac gcgcttcagc 65100
agagccgcga gcttctgttt taaaaatgga gaaagattta aaaattcaag ccatgagtcg 65160
ctaagtagca acgtttgtgt ggagggttct tctaacactt ctggtgacat tttactccaa 65220
actattgttt taagccaacc ctccataaat ccccgcgtct ggtaattact ttacgtatac 65280
gtgaaacttt tgtacttctg tctgagaaat ccaacagagg tggccgcgta taagcagttc 65340
cgctttcact accgtttatg tgaaagtctt caaactcggc taaaaactca tcaagactat 65400
attgcatata attcaacaat tgaagctctt ctctggggac actgtaattt aacttgtaca 65460
aaatccttat aaaaaacgcc ctaaggtgaa atccatttac acatatttct gtggtagagc 65520
gatcagcttt atatcgcaat gttgctgtac acagccctgg acataaagct aaatttgtta 65580
aatgaaaaga agcaatctca aacggctccg gaaatataaa gtcactaagt ggtattattt 65640
tacaacatgg atttactaca agcccggtac aacatgtaat taaaaaaaat ccacatgtaa 65700
taaataaact taaggtatac atacttctca tgttttatac ataatatgaa cgctaatgct 65760
gtacttatat aatatgtata gaattatgcc aatacagtca tggtaagtat atttaaggc 65820
ataatgcaaa catcaaataa tataacacaa aatgcacgct tccggatatc cgcccagctg 65880
ttaacccgaa aatacagaaa tgactacaca aacacacctg aaaccaattt tattctcaac 65940
atatgcaaca atatttcagg gtaacaccat gttaataaaa tacgcagcat gcacattttt 66000
agctaagctc ctcaaagaca atttcctctt caattgatga ctcatcaacc tcggtttccg 66060
tgttaatgtc ggaaatatat gagtccagga taccctcgtc atcgccacaa attaattcca 66120
gggcttgtaa aatatcatcc agtgaatcag ctgctaatgt caaggatgtt gtttctaatt 66180
ttagcccatc caaatggtga gtgggtgaca ctttgggaca agaccccgta aaattggcac 66240
gctccacaca aacggaagga tgcatttgat cgttcacgag gggcagcata tttattttta 66300
taacatcagg aaccggaccc gtcaatgacg cacatatgtc cattaagatt ggagtcgtga 66360
cactggttga attgacggtg aaaaaaaatg tattcgtgtc acatgcgtaa ctatttttta 66420
accacacgag ggcggaataa aacgggtccc cggcgctaag gccgtgaatg cggcaggcac 66480
ggctcattgc catttccaga tttgttaaat ctgttcgaac atacgagtct aaaatcaagt 66540
ataaccacga caagttcaaa caggaggcaa cgcggctggt ttccaccccgg acatccccgc 66600
tcacaaaccc gaactctcta ggaacgcaca tcaaagcgtt cagacaaaat tccgaaaccg 66660
aaaccctagt tctgagcacc atctggttat aaactgcggt taagagacga gcggccacac 66720
actgctgctt gtgaggattc agcttaggtg gcctgcaggc ttgctgacag gcccgtagct 66780
gcctggcggc gctcgcgcac ctttgacagc cgcacgccag ttccagagcg aggcagaagc 66840
gctctttgca gtcgcgccat atctcggcca taggagattc gctcgcgtga tcttttaggt 66900
gcatatatat ctgttgggta acaaacctca ctccctgtaa aaggggaata aggtccgttc 66960
tttctatcgt tttctctgaa tctatagtta catctttaag tataaccaaa gacgctaaaa 67020
acccaggatg gtcgacacct ttaaaaaatc gcaaaagcga ccggatggct ctatccgcgt 67080
ttccatgtcg ttcaaaggta gtcacgatgg atctccagtt agactctgtc tccctgttaa 67140
caccccttaac tggaatggaa acagccatga caccgtgaac ttcctgatgt ctcctaaaac 67200
taaccccgaa acagagctaa ataccaatga ctgtcacccc taccccaagc cacgcccccg 67260
tactattaga ccagggtgag taaccacgct atcctttaaa aacccatacg tggagtttgt 67320
aaggtaaaca gctcgtgtat atttcagacg cctgtcataa aatggatacc gacgacaatc 67380
aggtaattaa actttttta ttcaagattc aggagggcgt gtttacagtg gtgtaggtgg 67440
gagcatatct cgccaatggg aatggctgat gaatccacac ttagtgctcg gatcaggtgc 67500
tctgtgtact ttactagtgc cgatgaacct aaatttggcg tgtttaggtc gtaccaacca 67560
aaaaactgcc tggcctcgcc ttttagaagc tctatgcaag ctttaatcgt gtccaatagc 67620
ttgtctttaa aaatgcaact ctggtacatc tttatgacag tggtccaaaa aaaacaaaga 67680
tttaaaaaca cgttaaactg cgtgtcctgg tagtcctcgt atataacctc ctcaacaaga 67740
aaaaattttt taaccaaact cgccaggtac tgaaacgatg gtacggacag gtcgtgaaaa 67800
gtgtctatca tctccttcac ctcccctttg aaggttttgg ttacaccgac catgtgagat 67860
aggcaccagc tcagggggga ggtcggatcg tgagggggta acaattcgtg ggcgtggggg 67920

tactccagtt ttaggatctc cggcaggacg cgtacgagtt cttcgtcaaa tcctacgcag 67980
cccgcatcca ccaggggcag tactgagtta ataagtcccc gaattctatc gttcgtaatt 68040
ttgtggagct cctttagaca gtaaaccatg ttgctccgac actgcggtac gatgaaccgc 68100
tctccttgtc tgtgttgcat gagcacctgg agtatcttag acaggtatag ggcccgttcc 68160
actctaaggg ttgccgcttg gtagagcgga ttgcgggcat tgacggtcag tgtgttcatt 68220
tcaccgacta tcatgccgag tggcggatag tggaacttgt acatatgatt taaaagatgg 68280
tctttcccgt aatggcgact catggtcgtg ttttaacgtt cacagggtgt taaaaaactc 68340
agaaggtcct gcgtggacaa cttttgcgga ctctccgaac cgctcaggga taagctgtac 68400
cacgcggtgc ttcttataaa cgaagcatac gtcaaaaaac atgacccacg tgacctaaat 68460
aagctcacag aagaagtctg tcgatgtatt gtaatggaat gtgcctccct gggtcccatc 68520
tccgggctca tcgcggatct gaatctcttt aacctgttct gcctctaccg ggttcccgg 68580
gttaaaaccc gcggggccgc cacctgtaac gtcccgtgcg cagaatgcgc ccagggcatc 68640
gtgagaattc tcaccgagag ggccctttgt tgtaccgaaa agatgttcat agcgtctgcc 6800
tgcagcgggg tcgttatacc gccccagttg gccagagttc tgcacgacgt ctacgcggaa 68760
atgaaggcca agtgtttggg ggcgtggcgt cgtctcatat gttgcaggcg gcctatcatg 68820
gctattgccg actcggtcct cgtcacttat aataccctgg atgccgaggg aaaactggaa 68880
cttagactca aagcactgtg caaacttgtt tttcaaccca tctttcttca aagaatctta 68940
gcccctatgc agctactggc caacgggaag atggttcctg acaattattt taccatcacc 69000
ggtacggccg agaagaggcg ccctgtcgtg actggaagta ctagcgggat gacgtgtccg 69060
ggaagcagcc ttgtccccga ttccttaatc ctgccagtat gcgaaccggg gcttctcccg 69120
gcacccctgg ttgacctcag taatgtctta gaaaatccag aaatcatcct cagcgcccca 69180
cccctgagtc aatttgtcat cacaaacacg caccccagtc tgcctcagtc agtcagcatt 69240
attacgccaa cccagggcgt tgttcccggc caatgtttta tggacacgtg gaaagcggtg 69300
tcacagagca ttcaccacca ggcacagacg cctattttgg ccgccgcact aaccggttcg 69360
acatctgcgg cccctggccc gcatatcgca tgttccccag ttgccggcac gtctcggcag 69420
gtggaagggt ccgcgggcgt cgattgcggg aaaccagcat gcgttccgca gcccgcgtta 69480
ccgcccaatg tccccgccaa gaggatggaa acggtagcac agttgggaaa cgctcccgta 69540
aaaaacgtcc acatcggagg ccgcgtatac gctccactgg ttaatatacc aataatagac 69600
ttaacgtccc cgtcagggtc cggccagagt ccggccgata tcgccaacac tccagagtcc 69660
cgcatggcgg ccggctctcc gcccttcgcc gaaaccgccg caacggtccc cgctaagaga 69720
aagcagccac gcgaggacgt ggcagacaaa agactgaagg gcgacgttcg ggcgccgcca 69780
acagtaaacc acccttcccc gggaccgtcc gggatgcgcg ttcgcgagca gggcttattc 69840
gatttaatcg aaagctccac ggatgtaacc gcgaacgcat ctggacccaa aaacgacgac 69900
gacatgctag cggctatcct acaggacctg tatggactac agtccccccc ggccatcgat 69960
tccccctcca gcaactcgga caatgaggag atatttccag aggttagtcc gccatctagc 70020
ggccacggat cgccttgaaa gatgacccca agacgcaaga tcacaccacc gagacagtcc 70080
cacgcaaccc ggcggatgtc taaccgtaaa aatatcaccg gttggaaagt cccttggtgc 70140
gacaaaaatg agcttcaaac ttacccattt tctgtgctat ttcgggacac taaaaaattg 70200
tcgccagtgg cgacaatttt taagtctgtg taaatgtcac tctttctcaa gtttctgtcc 70260
tagaattttc acatccagtt tcatatcata taaagacacc tgttttgttt tcgtgaatct 70320
ggcgtgtttt tatcttgaga ggccccctaa aaacatgcct cgtgtgaaaa cacaacccaa 70380
gagaccccaa gtgcttgaat ttatgccatt agatctccac ggtggaacac acacggagat 70440
ggattctcaa aacctgtgtc ctgacggcca ggatctgctc gggtcttata tctatacgga 70500
gaataacggc ccgttttccc aaataatgca caatggacag agcaataccg ggacaggtga 70560
aagcttcggc agctacgctg ccggcgacgg ttttctgggc ggttctgtgt cagggatgta 70620
tggaaacaac accggagagg gcgcgtgttc taaaagaccg tccgcgtgcc gtaaacgctc 70680
ggctgcacta attcacgcgg cgtccgaggc gtctgtggcc gagcaaggca cctcacaggg 70740
ggcacatgcc gtatctgacc ggataggcag agacggtggc gctgacaata gactactcaa 70800
ggtgagtgcg cggctgtcgg acaaaacaaa gagcgccctt cgcagccatc cttgcttgcg 70860
ttgctattct ttgatgttta acacgtaatg tctgtttata ttaaccctgc catagcccga 70920
accgaacact gtgcgatgcg attataaaac gcaacactcg gacacgtcag atggacgcgc 70980
aaaacgcgcg cggcagtgca agacctcagg aaaaaataaa gcccaacagc tgcaacaggt 71040
acaataagca aatttagtaa tttgtatttt ttaagtttcc gaggttttcc cgctggcctg 71100
cggcaccgtg tccgtgcttt agttggaata cctctcctaa caagcacgta ttccaacagg 71160
agctggaaag taaaacaacg caactgaccg tcctcaccgc tcaacacgaa gcccttcaaa 71220
agcacgcatc atttttacag aagatgatga tattgatgtg caaaaacggg aacagtaaaa 71280
aaagcacgtc atgatttaaa gccacctcgg gttactagtt taaaactgcc taattgattc 71340
tattcaccct tcaataaaat aaaattttaa tacgtttgcg gtttgtccat ttcctgtctt 71400
aaaactatta aacaattcat ttgggaataa ggatttatgg gattaaggga ttaaattttt 71460
ttggatcatg ggatttggga acatacgtct gggatggagg ttatgcttca tggtctgggt 71520
ggcgtggatt gcacggggac ggtcggtgtg cccaacctgg cacctgacag atgggaaata 71580

cgaggcggta tacaggcact acctcgaaga gtgccgcaaa catgaaggct cggggagcct 71640
ggacggttcc ggacagacaa aggggtctgg aaccaaagca accaccgaag ctaatatatc 71700
gataagacct aacgttgtca catcaggtca aaataaagag ccgcctggga cagcaccgag 71760
ggccgaatca tcacacgacc tgccacgcat caagcaggtt aacgctctcc gattatcaac 71820
cccggaattg gcgcaaccac tcccggtagt aaaatcgact ccgcgcgagt cacagtcagg 71880
tgggacaccc tggaacgcgc gcccccacgc gttcattatg cacacaaacg acatgctcaa 71940
cccatctgtg gtcctgtctt tcagagccat ccgtgcgcgg tccacacgcg ataccgagca 72000
gtccgttcgc gatcggaaca cggtcacgac cagctatcgt accccctggcc gcccttccct 72060
ctttcaagcc agaccctcgt ctcacggtgc gcgtctaccc ccttcgcccc gaacgatggc 72120
aagatacgcc gagtcgcgaa caatatgcga ccaaaattga ccgcaaagaa taccacgccg 72180
tccgttcgac gcagtgcccg tatggacgct gcacgggacc tcaagcgtcc gtgtcagacc 72240
cattcctccc ttcgccgacg ccgtggcacc tggttcaccc cctgtgggct atggccggaa 72300
tattgggcct cacgatcgcc ttgtacacga tatatcaaat atatgcacat agaaactaca 72360
cgacgcttca ctgcgaataa aggtttattt attttgcaca ctagtccgcg tcgttatttc 72420
tggtcctgga ctgggcgcgc ctcctcctga gcgccccgtc cgcgctggca gcgattccgc 72480
ctctggtggt ctcgtccatt gacacgtcca cgcggatact aattccggcg agggcgtctt 72540
ccatctgccc gcgcgtcacg gccttggccg tggccgccct caccttagct tcaatcttac 72600
gggttgccac gttacataag gccgaaacgg cagctgcgat tttggcctcc ttttcggtgt 72660
tagagatgac ggggtcggag ccggacctgg tgggatcccc agaatttatc agctttttga 72720
gctgcttgtt ttcaacggac aatttttgca cctgagcggc taactcctcc attgtcagct 72780
ctttttggg ggcgcgcgtc ttgggacgcg tggaagacat ggcagctact acgcattaca 72840
gagtacagcg ggacgtttaa atactttcag gcggcccctg aggtcactcg tgcgccacgg 72900
tttcaacaaa tttaaaaaag cacatatatg ttaggtacag aaatgttgca accgcgacca 72960
gaacagagtt gatcaccgcc catatactcg aaaaactaga cagcacgggc tggtacgtgt 73020
ccgccgagca atcatagtca taaaaagagg gcgttacgga gggagccgcg gaggggtttg 73080
tagtcacagg aggcgtgccc gttcctgtag tcgaggaggt ggtgacggtg gcgcacaccg 73140
gcagacacag atataagcac gccaacaggg ctaacgcgag cacaattgac cctgtcattt 73200
tcaagccggc gttcggcagg caaccctcca ctttggcacg cgcgctctcc ttatatacta 73260
ccagttaaac attatggcag aagtgaccgc ccatacggtt ccgtacgcct ttgactcctg 73320
taagtttgaa ataataccaa aaaacaactc atctcgtata gcgttacgca acaaatttcc 73380
cgtcgtggtc aaaccgggag aacctctagt cgtgcccctg ggactaaaaa ttatccgcgc 73440
gccccagtgc gcattctttc tcagcggagc gccgacggat gaggtgtatt accacaccgg 73500
acttatagat caagggtatc gtggagaaat aaaactcatc gttctcaata agactaaaca 73560
ggtcgtcaca ctttaccggg gagaagtcaa cgtctcactc atcgcgttta tgtacgcctc 73620
gcccgggccc ctaaaatgcc cgatactgaa cctaccacat tactctctgg acgctggatt 73680
tgatgttacg tcgccacatg ccatgaccat acccccaacc gacagaaccc cgtttacgtt 73740
gagtttgtat tataagagtc cacagctcag cacccccccac gtcccactaa tcgtggggcg 73800
ctctggactg gcgacaaagg ggcttaccgt cgacgcgacc aaatggaccc agtcgttggt 73860
acatctgagg ttttataact ttaccaaaga gccaatagat ataccagcaa acagccgcat 73920
atgtcaggta gtgtttatcc acgaagacca cgtcccaagc gggtggaaca ttcttagatc 73980
ccgcgtacaa ctcggcagca ccctccagat atcgtgggcc aaaatcaggt ttaccgacgt 74040
ggccacgctc cccaaaaccc acccgctcaa ctcccgccac actcaaagcc aaaccgaacc 74100
ggagaccgcc cgcggcgcaa aggggttggg gtcgtcaggg ttataatttt aaaacatggc 74160
attattttaa agttgtttat ttttaataaa acaatccaaa tttacgttaa ccatcagctt 74220
gcgcatcctg tattattttg tctggcacga tattatccgc caccgccaac tggggcgtcg 74280
gttgtcgaat caccggatgg gaggggttgg aggacgcttc ggtgtctacc agcacattat 74340
taatttccgc agcccacgtg acagaatcat cggacgcgat cgtcgtagcc agtttttcca 74400
tttccaggag tgacacggtc tcgtcggccc gtttattcaa aagcgcgtga aaggcgctct 74460
cgtccaccgt aatatcccct atagccaggg acaacatggc gttccagata acgttctgta 74520
tagactttag cttagatatt tgaatcctaa ccgtctcctg gcaccccgcg ggtaccctag 74580
ccgcgtcaaa acgacgaaga tactcggtta ttttattgga ttgaacggcc aacagatacg 74640
tctgccgaag ggcggcggtg cacatctctg cctgcttcac atctcccata tccactcccg 74700
ggggcagacc catggctatt tcggtccgcc agtgagtagg acactcgtta agaggatctg 74760
gcaggggacg atattgacat cccacgcgtc caaacggact aaacccgcaa cacgtccaaa 74820
acgaaccagg cgcggccatt gccaaacaaa atggtagatg aaattagggc aattttctct 74880
actagtggag atatggccga agtaattacg gatatactga ctgaaacgca agcaacggcg 74940
tccttcttct gcgtgctcca cgatcggggc gacgcgccta taaatactcc acatgccgta 75000
attaaactct gcctgcccgc caagcgccca ggcggcgggc caaggtgttt accgttgatg 75060
gtgctgaacc taccggcgtg gcaggttaat ctattcctaa caggtgacgc accattgacc 75120
tcggataaca ttaaagaccg cattgacctg gctcagaccg aggaaatact cgaacccata 75180
ttaagcgtac tggcatgcaa acggtccgcg cagcagacca aacatgactc gtttaaatct 75240

aaggtggcct ggttcagggc aaagttcgtc tcggccctaa gaaaagtgta caaaatgacc 75300
ccatcccccct attggatgat aacgctgctg ggctcgttcg aggcatcgtt tgtactggcc 75360
ggcacgtttt atttttttcca gtctcatata tgcacggcgg agacgctggt gcatctaacg 75420
aggttattca gctccagtca gggccagagc ttggtcaccg taaacaccta tgacgaatta 75480
gggcgcgtat tcgggcggtc agatttcctt ggaattgttc caaatttttg ggcgtatcta 75540
aaatacaaaa tgcaacagga cgacgtggag tccagggcca tcgatcaaac aatcaactcc 75600
atacgtgggg ggttgatgct gtcacctcag gacctcgtac acttcatcta cctgtcattt 75660
tacgagtgca tgaacgcaca gacgttcctg tcgtactctc gtaccacgtc cagtttaccg 75720
acccccgcaa cggttaaccc tccgcagctg tgtcggcgct tagaagcgga ttttaaggag 75780
cacgtgatgg catattacaa taaagcaagc tacctgagta cttacataac cattttaacc 75840
gtacccgctc ctctcccgga cggatacgaa aactttcagg aactagcatg tcaatactgg 75900
tgtggacaat cgagagacgt ggcggaaatc atgactagaa ttaatgacca gtatccacag 75960
ctgaatctaa caaaagattt atccggtctc cttgacctag cggccctaga tcagtattcc 76020
ggaggcccca aggaaaacct ttttacggtg gcttcgagaa ttcccacata caggtgcgag 76080
tttttgaata aacagtattt tgttctaatg cacgcagact gtatagacgc gtactggaaa 76140
caaaacatta tcgtgcccga agacgcccaa ttgcaaggcc tgacggatca agatctcacc 76200
tccaggatat tttactgcga cctcggccta tctctaccta catttaaaca acaaattcta 76260
gtttcacgcc acgagtattt caacccacga ctcccagttt atagatgggt gttggatttt 76320
gatctaaagg taaccgaagg tagacggact ctaaacgata tctacaacat atgcgtaacg 76380
ctgcggcagg tcatattaga aacgttgcag ttaataggtc cactaaaacc gaaccacccc 76440
gtgtattttt ttaagtcggc ctgtccagcc gtaacctggc cggatgatat ctcagacacc 76500
gcgttctgtc actgtgacgc aaaaataggg atgagaattg taaccccgtt tccgagcgga 76560
tactgtttgg tggggtccgc tccactcgtg tccctgaccg acattttaaa ccgcgtggtt 76620
aaactggaca cgcggctagc ctcagagtat cctgggatct tagaagataa aggaccattt 76680
gactctggca tttacgccaa ggggcgatgc gtgcgcgtac ctcactgcta taaggttggc 76740
ccggggggcg agctgtcacg gctcctaaaa attattatct gtcaccccga agagtcagat 76800
aaatcggcgt atttgaaaaa cgcgtttaag gtttctaatc tgctacatca cgcccctgga 76860
gactctgtca ccaaaaacgg ccacctggtg tacgcgatca ctgacgaaaa cgagggcttt 76920
ctagaaagca aaaccaagaa taatcttccc aaaacaatca cggacctggc cgaaaaaatt 76980
gagcgaacca cagaaaaacc actaattgat tgggcagcga ccgcagtgtg gccaaaacta 77040
cacgatacca tacagcgttt ctttccggat gaccgcatcg gccaatttgc atctgtgagc 77100
ttcatgcact ccggagacaa cattatacaa gttaaaccac aaaaaggaaa caactttttc 77160
tgtattaatc ataagcaccg aaaccacact cagacagtcc gtgtattttt aaccctacat 77220
tccacaaaag aaagcgaggt cacggtgacc tttatgagtc agtgctttgc tgcaaaatgt 77280
aatcataata gcccaactgc acatttttca tttatggtac ccataaccgg cacgtaagag 77340
acacgaataa aatcatccgt gatgggccag ctctccggct actttgttaa ttttaacacc 77400
gcgtaacccc aaaaacaaca tctggacgag ttttgttcgc gaacgtgaac atggcacagg 77460
caatggtgag cctggaatac atgaaggata ttttggacgg taagaagacc ccttgtggct 77520
cttacgactc ccatttaaaa ccgcaactta tcaaacggct ctttttatac gatttacatg 77580
cggtatgtgt ttcacgcact aatttgtttt atagggggca tatcatcgtc ggacttcgac 77640
gactcgtcct cggacgaaat ggacgacctg tccccaacgc cggagccaga accgtccaca 77700
acgccgaaca gctttccgga ggggcccaaa tcacaagtgg tggcgttacc taaaatccgt 77760
aaaagatctc gatctgagac gccggtaaaa attgagcaca gatctccact taaccgctcg 77820
cgatctcgat ccagaacgcg gtccgggtcc ggtcaacgat caaaccagtc aggcagatac 77880
gtcaagagat tcaaaccaac ggttgatgca ccccgtcatc gagaaccgtg gcacaggggc 77940
ggtaagggaa aggccccgtt tatccgcaga gacgcaatgg ctggccgcgg tcgacgcaca 78000
tacggccacg actatcgcgg aaaagccgct ttaacgcgga gcattaaaga gtctattaaa 78060
aagatgcacc ttccatccac catgctctct cgtgcgcacg ataaaaaggt attcgaggga 78120
ctgttgccac gacacctggg acagtgcttt caggtgtgcc tgccggcgcc gccaccgctg 78180
caacccgagg tgttcacgga tcgacagctt accgctatag tcaagtctgg cgggcgcaga 78240
gacgcgctgg tggccaaaaa agttagcctg gctaaactaa caagcctata caaacccctg 78300
ctcacgtttg tgacggggag aaacaaccag gcccactggt tggcgacgcg caaaaacacg 78360
ctagcgtctg cgggactcga ggctctcgcg gcctttatcg aggaaggtct ggcgtgggcc 78420
caggtgtgcg tatctcaaaa cagatcgttg aacgacagca acctagatat cattctagac 78480
agtagtcaga gcgtgtgcac ctggttcatc tccaaaatca ggcatttgca catacagtgc 78540
ttcttggaga atcagggtga ggttagcctg gtgaagcagc taacttacct cgtgtgtata 78600
aataaccgcc tggcggaggc ggcgaacctg gcgggagagg tgaaactgaa ttttaaactc 78660
ggaatgctga ttgggttcgc cctaaccttg ccggcgctcc ttgcagagca taagttgtct 78720
ggcgagagcc tgtacctgtt taggtctttc ctggagaaat atagaccggg agatgtcatg 78780
ggactcttaa attcaatagt ggtggagcac tacaccaaat gtcgtagcgc ggagtgtgtc 78840
attacaaccc acgccatggt aggatctggc gaaaacaaca agggactttt tttctttcca 78900

gtgtaatata taacccatgt gtgtaagtag ttacggtata ttattcacgg gcgtttaaat 78960
gcaataaccc acataacaaa ataaaaatgt gtaaaaccaa acacgcgtca agctttttct 79020
gcaagggttc tgtgctcgga acaaataggc agagttcgcc acctagcggc acgttctcta 79080
tagccttggt atacagcgga tgttatttac ggaagtgcca catgaaaccg tggccgggtg 79140
cgtaacagta aaccgggtcc aatcgcgcag tttacacaac acataaaata ggcgaccgtt 79200
tatagcggtt tatttgaaca tacaacacac atttcaggag tttcggtctt tataaaaatc 79260
ttcattcatt caaagtgccg atatatttca cgctggtgga caacaataca tatcggcgcc 79320
tgtgattgat cctcctcctc gtatagggac caggcagaca cccaaatctg tgcaaacgga 79380
tcgggtggcc cgggatcaga gggtcgtggt gactgggcta actttaacat gtattcatcg 79440
gtatcaaaaa cgcgatgtcc caccccagat ggcacgcgtc ttggttcaga ggtgtttgag 79500
gtattgccac ggaaccataa attaagaaat ccatagccat tcatcataat cccatactga 79560
caactgcaca aaacaatgcc gtcctccagg gcgtgtaacg cgttcacaag tctgctcaac 79620
ataaaatcgt catccatata aggcgccggg ggcaaccaca actgaagcgg cccgggagcc 79680
atgcagacgt ggtgttcatt tttatccgtt ggtctcgtgg ataaacgaac gcccatccca 79740
gaatgggtaa aagtctgtgc ttgaaggacc ccataataat aaacctgaat atgcataaat 79800
ccccacgggt cgtcaacaga tggcgccaaa cacgcagaaa gaccgctcga tgcagatagc 79860
ggcagggcgt gttccccggg caacgcaacc gcgcccccctt gactattgtc ctgcctaccg 79920
tcttcaccgg aggcggttgc ggttgcagac gcctcggtgg ccggggcggt catggtgaag 79980
gtaccggcgg gacgggtacg aagtagcccg agcagctgct ttaccgtgta ttcagatacc 80040
ctgcccgcct cgtcggcgcc cttgcgctta aacatccgaa acctcaagcc ccgcagttct 80100
gccgctaccc cagaggcccg agggcacaac ctgcataaca tctcttcatc ggagcgcaac 80160
ctgaagacca cgtttgcaaa tgaaaaactc cgtctcaaaa cgtccttttc ttctatgaaa 80220
taggcactct ttcttatcgc acccaatagg cgcgccttgg cctggcccag cctagacctc 80280
ccggacggaa gggggatgtt tctctggttg cagtagtcct cgtacatctt cccatcccga 80340
aggtaatcaa atcccacaga ccacggggtc gtagtcgcca gcacgaccct ggtttttcg 80400
tcgtctgccc aaaacagtcc cgaatacgtt gcaagatttg cgtgagtaac gatccattct 80460
cgaagtccca tggcggacgg accaccgctc tgaggtgggc gccaggtagc catggctgta 80520
tatgccaggt tgatactcac cgctttcagt tcgtgttaaa ataaccaaac tccgcccatg 80580
cggctaccgc cgtgggcgca gatgtgtgca gccgtcacaa aagaatccca cctgcggcaa 80640
ctaattccct cagcgcccag tcagctggct accggttcat ttcactcacg agcacgccac 80700
cccatcatct tcaccgctcg ttcccgcgtc cggatcgtcc attgaatctg gtccggagtc 80760
gcccagctct actgccgggt catgcgattg gctcgtcccc tctcctggat cttccatctg 80820
gaatgaacag tcggtgcgac cctctgatgt tacagtcgtc ccggggtaaa aacaaccctc 80880
tgtaccgcat acagacgagg tttcgcacac gtaacgcagt ctcatggcga tgggcgcatc 80940
gacgggattc ggcgactgga ccgacagaca cagggcgatc tgtgcataat ctctaggtgg 81000
ttcgtctcca gggttcgggg accttgcaag agctctaaga taatccacca tatcaaaaac 81060
ctgaacgggt tgcctggatg gcaggtcctt atactcctcc tcattgttag taagcacctt 81120
catttcccgg ttctccaggt tcttcacgca aatgccagta tccgtaaggg taatcaacac 81180
acccttctcg cactttttgt gcataaccct tagatcttct ctggcaaact gacagtcaac 81240
gacccccggt gacggtaaaa acaaagtcct ctcatcttgg aaacaacacg cgtggccagc 81300
gaggacggga catggtcgac tgcagacccg tattccagca tgaacacgtt cgaggccaac 81360
ggtctctcca aaatatagaa acgttaattg gatggcccct ggcagccgat gctcgcgcac 81420
acgcttgttt ttgtccatcc gactccgagc caataccgtt gcggtaaaaa cccgtcttct 81480
ttcacgtcca gacccacccc cctggtcgta ttggaactca ctgattaaac caaggcactg 81540
agtttcggtg gaagacatca gattacacat tgcacaactt accaccggtt ccggcaacag 81600
gcgaaacagt ctatatcgag cacccggtcg cccataggcc ttattaataa atttaagctc 81660
tcgtatactt ttgtggtgtc tcatgtcgta caagaatctc attttcctga atctgccata 81720
gtgggacggt ttttcgccct gacatatgcc tcgcatgttg cagtagtcat caaaaaactt 81780
ctcgctattc catacggacc gatctgcagt tactttgttc cacgggaccc ggacaacggt 81840
cttgtcctca ttctcccaaa ccaatccgcg gtactcgccc gattcgaccg cggcaacgag 81900
ccaggccctg atatcgactc cacggcccgc cattccgcgt attgggagct tgacggatca 81960
gctctcaaga caaatgaaac ccttgcgttc acggcaccag ttcataaatg tatctaaatg 82020
acggagtcat gtgccctaaa catatggcta agactactca aactgtgctt aggggaagag 82080
tgacaaacac ctggcactca ataaaatttt gaggtcgggc attgacccgt gccttcctct 82140
ggttagtatt aaccccatcg ccaccatatt caaggaacaa aacactccag atacaaaaca 82200
ctttattaca gatagaaggc gctcaccggc ctcacaaaac cggtcgcgcg cacaggtaac 82260
attaaacatt agcggcggac aacgcccgcc acagacactc gtgccacagc tgaataatta 82320
gcggtacagt gctttgcact ccgccctcgg gtgtaggtac gcccccaaaa tataggtaaa 82380
cacaagcttg tggaagcgcc gctccgtccg gagatgttct ggccatggct gtcatatagt 82440
caaccaagtt aaatgcacga cacgcgttgg cgcacgacag gaccgcaacc gcccccgtag 82500
attcgacgtt cccttgaacg aacgccaggg aatgccccaa ccatctaatg taaatgccgg 82560

[redacted]

aggggtcggc ccaaaacgcc aacccgcgca aaagatcttt atccaggaag cgcaacgcct 82620
cggcctgaaa ctctgctaag tggcctggaa tctccggaac cagtgcctgc tcggcgatcg 82680
gtgcggcgca tatgtgccct tgagggcgac gttcggacaa tgggagcaca cggatgccgt 82740
tggacgattc agtcaccact tcaccaactc tgtccccata atagtaaaac attatacgta 82800
gcatccaaca ctgttcaccc ccggctccca ccgcaccagt gtaccgaaca ccggcgccaa 82860
ccggaccaaa cttattcacg gcctcctcgc gcaacccccg taacatagca gtagtgccgg 82920
aggcctgatc gcaacgagca cacgacctca cggtaattgg taaaaggcga atgattcggc 82980
atctgtccgg atatggagtc gaaaggttgt ccttagttga cacgtccgac acggtttgac 83040
tatgccgtat ggcgctggat agccagtttt tgcattcctt agcggttaat tccctgccag 83100
ccgcgtgcaa gatgcccctg gatcgacagt agtcaataaa aatgtttttt tccgcctcct 83160
cgactcctct gctccctctg tcgtgattcc acggtatgcg gattaaagtt ctttcctcgt 83220
caacccaccg cattccggga tgacgacctg tctcgcagca ttccaccagc catcctcgaa 83280
ggctggatgg tttggtcact cggaccggcc gctccatatc ctcgccgcgt gcacggcaat 83340
gccaaacctt gtccaagtta cctcctgaca agcacctgcg accctcacag acctacgcgt 83400
ggcaaacggg aacttgtggt ctagtctgct gggacagtac cttagcgtta tttttatctg 83460
tcgtgcactc ctgcggcgga agtttgtggt ttatgagcac caccgggaag gattcatgtg 83520
tccctgcacc agctcgtcct gacagaccac cacagaaata ggaacttgtg ggctctcttg 83580
ggcaccaggc tgctgtgaga gatacagtgt tgccccgttg cacggcgcat ggctctctgg 83640
tcccggagag cgggcaagct ttctcatgaa acccatgaag ttaaaaattt tctgtttaga 83700
gaggaacatg acctctcgtt ggataatatc accattacag tcgttacccg ccgcgtacag 83760
cgtttccctg ttcctattat cgacgtacac tccggtttcg ttgccgtgta cgtagatgcc 83820
ctcgttatac gcgaccagca tcgtttttat ttgcgttatc tgttctgcgg tcagtttttc 83880
acgatctgga gatggaaacc agacagtgaa cccggtgcca tagaaacaca gatgaccggg 83940
atggggaacc gggcgcggat gcaaacgaac acctctccga tcttgggacg tgtgatccag 84000
tgccagaacg ccaaaataaa acacctgaat tctcaaccgc caaaacggca caatttgatg 84060
aggcagcggc gcttctggga ccccgagacg atgtgctcgg gcgtacgctg atctcgtcgg 84120
agggaccaac gtgggcaagg gggcgggcag cggcattccc ggctctaacc cgatgccaca 84180
cacctgctcc aaataattta taatcccttc agagttctca tcaaggatac acgcatagca 84240
tccagaatcc atcaaaggtt ttacccggaa ggcggtccag ccgccaccgt gcgtccgcac 84300
gaccccatgc tccacttctt gaaatccagc attctccctc aaacccctca aaatcgtct 84360
cttacactct agcaggtttc catctagccc aacctgcaga cctctctcca cgcagtaggc 84420
gacaacggct ggagttccag ctgcccccctt cttccaagac agcttaaagc tttttttccc 84480
tcatcgcacc aagttaggtc tgaataggtt ttctcgttga gatggtaaag ggtccacgat 84540
cgtagatggc tgtagttaga atctcgattc gccatgacag cgtacaccta tacaataagg 84600
cgcgatacga gtgtactttt cccctgatga tatttgcagg gctctatctc ctcaaaacgt 84660
gcacctggct aagcaagcag ctctatttct gccacttcct catttataaa accatattgt 84720
ttctggtctc acgttttgca gctcacctgc aacgacacgg tgaatgatcc cgagtttcgc 84780
tcctccagtg ccttgacaaa cactggccca gagttaaaaa gtgaagcaaa aggcatagct 84840
ttcgatttca gtcaccggca gtgagtggtg gacattagaa agtgtttggc ccacacattc 84900
agctgtgcac tcaagcccga ccaatatggc acgtgaccat tcccctatca caacactagt 84960
acagaaaaaa cgaaaccaca cgcgaggtga cattttccag gttagagaaa tttattaagc 85020
cggcgaatcc atctaacaaa acagcaaatg tttattcaaa gtgcctatag atttctgttt 85080
gataaacaat aatagatagt ggagccctcg aagagtttcg ctcctcgccc agggaacaac 85140
ccgaaaccca gagtttaaca aaggccgcgg ggggacccgt gtcgcgcggg gagggggacc 85200
tggccatatc aagcaaataa cgatcagtgt caaatgccca cacgtgcatc ccggtgtagg 85260
cacgatgtgg caaactgtac ggttccacgg tatttccctg aaaccacacg ttaggacccc 85320
cggtaatcgt gaacataatt cccctgctgg tgctgcagaa aattattcca tccacgagcc 85380
atttcaaagc gtcctcgagt ctctgcacca cagacgactc agtcgtctgt ggtatcggag 85440
gaaaccacgc ctgaaggggt ccatccaaaa cacacacgtg gtgtcccgcc ttatttgttg 85500
gttttgaaga tagcctgatt ccctgcccgg tggctgtcaa cagctcagcc tgcaattgcc 85560
catagtaata aatctttata tgcatgtgac cccatggatc gccggcagat ggcgcagcac 85620
acgccataag tctggcgcgc agaaggcccg aagaatacgc tgcatcgtcg tcctgcgcgt 85680
cgtgtccatc ttgatgacct tcccgatacc gtctctcttg atcacttacg gacggtcccg 85740
aaagggcgtg aagcattccg cgcccgagca tggcaaaccg catattgcgc aactcaaggc 85800
ggagcgcgca cacgcgcgga cacaccggac aagacatctc ttcggatgaa cgcaacctaa 85860
agaccagcgc cgtgaacgga aattgatcgg tagggaattc cttatcttcc tcaaagtact 85920
tgctctttct tacggctccc agcaggcgcc ccctggcttg acacagcctg ctccttccag 85980
acggcagtgg gatgtgtctt aactcacagt aggcatcata gtgttgaccg tcgcgttgat 86040
aatcataatt gggaaacgac ggtggtgtag ctgccaacac cagcctcgtc ttttgctcgt 86100
ctgcccaaaa caggcccggg tacagaccac tatcggcttt gaacgtcagc cattccttga 86160
gggccatgcc ataacgagat gggccccttc tcgattcccc agcggccatg attcagtcgg 86220

caacgtttaa aacccatgta aagtttcagt tcaatttaaa atgacacaac tccgcccatg 86280
ctgtgacaaa cataaagaaa gtggcacacg tgtgcgcttt ggccgtcctg tcagatgaac 86340
cccgccggta acatttactt cctcccatac ggtgcgaggg tagatacaga tgtgcttccc 86400
gttgttcccg aagcaccgac acccgctcca gacggtccct cgtccccgct gtcatctccc 86460
gcggcaccgt tgctgacatc atctgacaca tcatcattgc acacgtagcg aagtttcagg 86520
gcaatgggag catcccatgt atttcggac tgagtagacg ggcacatgaa aatccacgca 86580
cagtcacgcg gcgggacatc tcctggagtt ggcgatctcg ccagctccct caaataatgc 86640
aagagatcga acgcctggag cggctgcgaa ggccgcagct catagtaatt ttcagaatag 86700
ttggtcaatg ccgtcatgtt gcgttcttcc aggtttttca cacagattcc agactccgtc 86760
ataacacaga tcaaacctct ctcgcatgtt ttctgcaggg cctgcaaatc tgccctggca 86820
aactcgcagg gaatcaactg ggacgacggc aagaaacacg tcaatggatt atcggcacaa 86880
cacaggtgtc cgggacgttt gggatcagga aggccacata tacgcacacc cgagccggct 86940
cgcaaaattt gaactccgac gttctctcca aagtaaaaaa aggatacctc aacagcccct 87000
agcacaggtt gcgcagaaga gtcctggagt cgtttggatt tccgtttata aacggagact 87060
ctcagtggcg ccttagccct cggccgccgt gtcctctcga aaacatcctc tcgaacgcaa 87120
ctgcccatgg aacagctatg caacgatgtg gtcgcgatga ggttacacag atcgcagccc 87180
accgtaacct caggaagcaa ttgaaagagt ctgtatcttt ccccagagac gcccccggcc 87240
ttgctcggcg gtgtcagctc cctcaaatac ctgtgagatc tcactgcata taacatccta 87300
atttcttaa atttctgggc atgtgaaccg gtttgacaca cgccacgcgc tgagcaaaaa 87360
tcgtcaaata tcttgttata atcttcatcg acgcgactgt cggtgcaccg gttccacggg 87420
acacgtataa tggtcttgtc ctcgtcatcc cattccactc cggggtatcg cttagactcc 87480
acggcctcga tgaaccaagc ttttaaatcc atatcgcgtt ccgccatctc gcgtattcaa 87540
tacctctgac cgagcggctc agttaaaata ccgagccccc agaaccccag gcctccgtaa 87600
caaaaaggat ctagatctga aagaagccta actatatacg cccacgggaa aactctgtaa 87660
caaatgaaaa aaatgttctc agaaaacgaa gagaggaaaa ctgacgacac gcaatctagt 87720
caaaccagaa cctaaaggcc gtcactgacc catatctttc cccttcccga gcgagtgcgg 87780
tttctcacat gccgtgagca aacgaaaacg ctacacataa taagacacgt gtgagaggaa 87840
aactttattg cagggacagg gcaaaagcaa gctgtgcacg gtaacagtat gtgtcactgg 87900
gccgcatccc cgcacgcccg ccacagacac tcgtgccaaa tttgaatgga caggggggatt 87960
gtgggacacg gctcgccttc cggagtggga acgcctccta ggtggagaga cgcataggcg 88020
gctggaaatc gcgttccgtg tggcggggtg gctgacaccg cactcagata ttgggcgagc 88080
gaaaacacgc gcagaactcg cgcgttgtta attatatcca ccggctccag cgaacctccg 88140
ttacccagca ggaaggctgg aacgtgaccc atataccgaa tgtaaatccc agacccgtct 88200
gcgtaaatca tcaacccgcg tatgatctcc ttccccagcc actttagcat gctcgtctgg 88260
tgaggaaaca tatcctgtgg ggtatgcggc aacagggcct gttccggccg aatgcccgca 88320
caaacgtgtc cttgcagtgg tcgctttggc gccggcaaca ggcgaacgcc gttgggcgac 88380
ccgacctcca tgttcccaac caaccgccca tagtaataaa caaccagccg tagcatccaa 88440
cactgctccc cgtcgcctcc cacggcgcct tcgtactcca tcccggcacc gagataccgg 88500
aatctggcgg ttacctcgtt gcgtacgtcc aggagcatgc cgccggtggc atccgcttcg 88560
ttacagagcg ggcaagatcg tacaaagatc ggcaacaggc gaataacgcg gcacctgtta 88620
ggcgccggcg ccgacaggtt ctcctcagtg gatacatctt cgacggtctg actgtggcgt 88680
atagcgctgg ccaaccagtt cttacattcc ctgacgctca atcgtctgtt accggtctgc 88740
cgtatgcctt taaattgaca atagtccaga tacatgtttc tctcatcgtc agagacgcca 88800
cccgcgccct ttaaatgatt ccaaggtaaa cgaaccagag tcttttcctc atccagccaa 88860
tgcattccag gatgcttatc attatcgcaa caatcgagca accatgccct caggccagac 88920
ggccggttta cgcggatcga ccctgcccgt ccttccgcca tctctgaggg tccgctcgca 88980
taaactggct taaaaatcta tcccccccg tgcacctgta ctttacagtc acaccttcgg 89040
gcacacgccg tagggctggc agggacggca cccagcagct attttcatct ctagcgccgc 89100
ctactagcgg aaatctgcgg tctacaaaat ccaccgagcc gcccgaatgt caccatgagt 89160
caattcatcc tgacagatca ccaaagtgac agggacctgg ggtcttttac tggagccagg 89220
gcgaccccca aggtaaagag tgactgcgtt gctgggaaca gtgtcgccgg gaacgggtga 89280
ccgggcgagt ctgcgcagta gacccataac atagaagatt tgagatttga cagaggatcg 89340
cacagcccgt tgtgcgtggt tcccctctgc gtcgttccct gcgcagcgaa ccgtgtgcct 89400
ggttctgtta tccacataca tccccagctc attgccatgc aaaaagacgc cctcctcaca 89460
cccgtttata atctcacaca cggtttctct ctgctccggc gttagcggcc catccaactg 89520
aggttttgga agccacatct tatacgctgt cccatagaaa cactcgtgac ccggtttggg 89580
gtcttggcgc ttatgcaatc gtacccctct acggtcctgg gacgtgtgtt ccgctaccag 89640
cgacccaaaa taaaaaactc gtatacgtag ccgccaaaat ggagtgatct ggtctggatg 89700
cggtggtgca accgtcgcca acctggcggc cctggcatat acggaccgcg ccgcaccgtc 89760
agcctccgcc ggagcgggag gaggcagggg cgtccctggc tcaagcccat ggacgcaggc 89820
cgcctctaag taatttagga cggtctccga atgactatct aagatgcagg cataacacag 89880

cggatccacg gcgggtctca agtaaaaaaa tctttctccg ccaaaacgcg ttgttgttgc 89940
ccgggcatcg tcctgttcaa acccagcgtt ttcccgaagc acccttagca gtcttttccg 90000
gcactccgac acgttgtgct ggcgtccaca ttccagatcc ctatccaggc agtacgcaac 90060
cacaggttgc atgccactca tacctcggtg ccatgatatc ctaaagctcc tcttttcctc 90120
gtcgcaccaa cacaagtgtt cgggaaacgt gttcgcttcc aaattggaaa taatccatct 90180
tcgcaggtga ttgtgagtga tttctatctc ggtcattgtg aacttggtcc acgaaaacta 90240
actcaaaagc cctcaccact cccctaaact gtagctgaat tttatgtcca ccttataggg 90300
tgcacgtgta cgtatatgcc ccgtgaaaac ctcgcacttc ctcataaatt atccccaaag 90360
cgtttcgggt cgctcaccta aactaactcg agcattacac ggttcatagt gtcccgggtt 90420
tcgccactcc cccaccgcga cagagaaacg aaacgttcaa tttaaagggc accaagctca 90480
actttattta aggagaaaaa aaccaccgcc atcagaaggt ttccgcaggc acacaccccc 90540
aagactattc ttggcgagtt cgcggcacgc tgtagttttt ttctgattat ttgtaattcg 90600
catccaatcg ccccaaacat agtcaataaa aaaaccgtaa acacacacac tctggtgaac 90660
aaattatgcg ttcctggccc caagacgtgc gtgagcgtca gcagcacagc ggctgctaaa 90720
acataaacca gggctaaaac gtttgaccga aacacgagtc cgaggcaaca aaaaagccca 90780
acgtggaaga caaaaaaata gtgtacaagt ccgaggagca gaggcgtcgt caggccaaca 90840
gacagcgtac ccgcggccgc ccctccaatc aaaacagtca cgagataaaa gtcccatcca 90900
cacacctcaa gcgcttgctt tatagagtat gtgatgtacc gcctagaggt gagcgaaaaa 90960
attgcatggc gtcgctgcag gcccgcctca taaagcggat gtgaccgcag cgctctgagt 91020
gaacaaacgc cacctactat aaaggccacc aggggcagca aaaacaccgg ggctaaaaaa 91080
aggttcttaa gaactaaaaa ataatacacc gagaacgctg ctaaaaatcc aagttcgtag 91140
tatctgtggg caatcgtggg gcacaaatat accacgtcat tcgccgcaac aacacagagc 91200
agcaggcaca ggcacgaccc cacaaaaatg gaaacgtgaa cccatggaag gtcaacgccg 91260
agggtggatg ccatagaacc ggcggtaata agaaacgccg agatggtgca cggtacatcc 91320
acctgccgcg agagccaggc cggtagccca aatatacaga atagcaaaag cagcccatgt 91380
aagccaaccg cgccccagga gaatataatg ttctcaacgc taatcagggg ctctacaaaa 91440
tacagcgagt gataaataaa acacagaata aacggggaag ccgagaccgt cccgctcata 91500
aaagatagcc aggcgaggga cgcctccgac gtgtacgtgc ccatggcgcg gcgtttaaaa 91560
caacgggttg aacgtcgcgc gaggcgtctt ggtttttcct ttcccgcgac tgggtgagct 91620
gctgtgtttt cgcttgccgt gttttttttc cggtcgagag ccgctcgtcc gcagtgattt 91680
ggcgcccccg tctgtcggct cggcaccgtc ctcggcaatt gggacgcact gctcttttaa 91740
gtccaggtca ggcgttctcg gaggggacgg gagctcagga acccccctcgg gatcggcacc 91800
ttccgagaga ttgcccggga cggcattatt ccggtccgtc ggatctacct ttaatagcat 91860
gacgccccaac ctgacgtcag caaggctatc tgtcagcggc aggccctcaa cttccagaat 91920
ctcattggcg tgccagcgga aaccaggcaa aaaaaatccc ggaatcttac atgcgttcaa 91980
tgcggccacc agagactcta aactcacatg agtcgtcagg tcgctagaaa ccacacccac 92040
gtccagcttt cttcctctag cgcaagtcca cttggcgtcc ggggtgtcgg gtttaaaatc 92100
aataatttta ctaaccccgt cgaccgtaaa cgtggcaact gccagcgttt cgcttaacga 92160
gacggtgacc acctgcttgg ccctggtttt ctgctgcctt agccacttta acaacaaagt 92220
cagcgtctta accgtaagga acaggcgatt ccccgtcccg gcctgctcta gcctgtcgtg 92280
gaacgagggt aggttgggtt cgtccacgac gctcttgtgc acggtgctcg tgtgattgtc 92340
cccgtaggcg atggtcgtcc taacaaactg cggctttcct tccgcgtccg actcgtcccc 92400
gtggtggtag aacgccacgt tcatttccgc aatgttaact ccaaacatct cacgggtgtg 92460
tgtgaacgtg ttcccaaggg aggtgtttcg aaagctaaag ctaacctcgt ccgtcatgtc 92520
agaatatacc atcagcggcg ccagtgcatt tttaagcctg agaccgagaa cgcccgcgtc 92580
ccccacgctg gagagcacgc taaggacggg ggccctgccc tggccgctga tctgtataac 92640
gcctttctta acgatgccct taatgtggtc gtaaaccctg ctaatgctgc caagcgcatc 92700
gacgtccacc ctggcgccgt aatgaaacga gacaggcatg gctaaggcga aattaacacc 92761
ctaaaagacg agaagagaat ggcaaacacg acagacacag cgttcaatgt tcaacctttt 92820
aatagagcct ctggcccgga aaatgagtaa catgggacac atggggacta cagatcatta 92880
tgaacggatg tgacatattc agaattatca cgctcaaaaa agttcacact cttggcactc 92940
gtatacgaca agggacacgc gtttgggggg cgggcaccga acaggggcgg tatgtttatt 93000
gactttagaa tcctgtcggc cgtggcctgc aaaaactgct cgatgtctcc gacgttcacc 93060
atggtcacgt ctcggctctt ggaacgaata aacagagtct ctatctccac ggcttcgcga 93120
aagaggccgt ggatgtatgc gacggagggg gattcgtctc gggagatcat ggtgttgtac 93180
agcaagctgg cggccctggt gtgtagcagc tcgtccctgg agatgtagtc gttagccagg 93240
cacaccccgc gcatgattcc gcgaaggcga aaaagtccta tagagtaaaa tgaacttata 93300
aagtatatcc cttcgataac caggaacagc agcaccttct cggcccgcgt ggtggccttg 93360
cgcactcggg aatgcagcca gtcaatcttc ttcgccaggg tagcgtcctc gactattttt 93420
ttggcataaa tcattaagtc cccccgatta ccgccaaaaa ataagttcag tatgttggca 93480
tacactttgc cgtggatgtt ttccatggct acctgttcgg cgtaatagtg ggtcacgtcg 93540

tgattgctga actctttaac taggtcttcg atgttaaagt ttaccagact ctctgccata 93600
cccaaaaatg tgaagagaaa ctggtaaaac tccctatcgg tgggactcaa tagtcgcacc 93660
tcatcggagt cctcgtgcag ggaaatttga ctggggaacc atctgttctg ccatgtctcc 93720
tgggtgagtg ctaaaaagcc agggtggtcg ctcgcgtaca gaaaccgatc tggggactcc 93780
atggcggcgg ttactattga caggccaaac ataccttggt ccgctcaccg tcggcctttt 93840
gtacgcagtt agactccgcc ccctctgaag agatgttgtc acccacgacc tgatctaagc 93900
actgcagggc attaagtttt gtctgtttct gaattcgaca ataatacagc atggttttta 93960
accctagcct gtatccggtc aggagcaagt ctcgaatata actcgcccgc gtggcctgct 94020
cctcctttaa aaagaaatta agcgactgac tctggtctac aaatggagcc cgcgcaccgg 94080
cccgctccaa ctgcttaatt gggcagtaat caaaggccgt caaaaaaatc ttgtatctgt 94140
ccttcagggg ttctggaaat gaagccacgt cacctccgta acggcgcacc gttcgcaggt 94200
caccaggctt cactctctta aaaaaagtca cgtttggctt caagatctct tctttgctgg 94260
tgaccttcga tgcgatattg gcaaagaacg gataaaatgc ttcggtgtac cccgtaagct 94320
gagaggttcc ggccgttggc atcagcgcca aaaactggct gttgaaaatt ccatgctggg 94380
caatgctgcg ccccagctgc tcccatctct ccaagggagg gtgggacggc ttaacgccgt 94440
cccacgtttg ccagtgaaac acaccctgag ccaatcgact ccgctcccat ccacgaaacg 94500
gagtcccctc gccgagtaaa acaatctcat gactggtttg caccgccgta aaatacattg 94560
cctgaaaaat ctccacgtcc agcttggcgc tctcggcgtc aaggtagcca aagcccagtt 94620
cggcaaacac gtcggccagt ccttggacgc caatgcccat tgatcgttcc tcctggccgc 94680
gcctaacgct ctcggtaggc gccgtcccgc ccaaaatgca cgcattgaca atgataactg 94740
cggcctccac ggcatcatcg agcagttcaa aaccaaatgt tacatccccc ttcccaggct 94800
ccacaccaga ctctcccctg tggggtctga ggcactttgg taggctaatg tttgccaagt 94860
tacacacgga agcctgaccc tcgggttgct gcacgatttc cgcacagaga ttagaacagt 94920
ttatggcgct gccttgcgtc tcacaccagt ggtgtttgtt gagcgcctcc tttagcagga 94980
cgtagggact gccggtctta atgacagtgt taataagggc atacatcatc gattttaacg 95040
gcaacgagct agagtgtttg ccagcggcca ctagtctgtt gtattcaatc tcaaattcgg 95100
caccgtagag ttttaggaga tttggcgcca cctctggcgc aaacaggtgc cactggccat 95160
ctgggtttgt ttcgtacagt cggaaaaaaa gctccggcac acacacgccc tgaaacaggt 95220
tgtgacaccg ctcctgattc tccggcatct tcgcgttcaa aaaatcacaa atctgatgat 95280
gccatagttc catgtaggcg ctggcaccaa caggacggat attgttatcg ttgaaatacc 95340
caacgtgggc gtttattaat tttaaacagc tggtgatgtt cttgtgttcc gcaaaagacg 95400
agacatctat ccccacgcct gacttgctag cgagaagggg ggacatttcc tcatgaagtg 95460
ctttgagggt tttgtcttcg gtcgccatgg acggctttaa aataaaacag ctagaaagct 95520
gacctccgcg aagcccggct gaccttaata caggcgttgc gcagcacaca atctgtgacg 95580
agatgtaatg gaacgcgtaa ccaaccaggt acatctcatc cagctccgtt tcgctctcca 95640
ccaggtgtct gagggtctcc cgcaaacacg gaaattttat acactgacag gccacaaaaa 95700
cagccaccct cataaacatc tgggccacgc tttcaaaaat gggtgaagaa ccctgggttc 95760
tcagcacgta cgtatcgtaa aacctaacgg cagacaggta gccgcagtta acgaaatttg 95820
tgtatgcctt gctctgctta aagtcctgta aaagaccgtc aagggccgct tcgtgttttg 95880
acataaacgc gcgaacctcg tcgcttagcc tttccccgaa taccgcgaga tagtctcgca 95940
ccgtaaccag gcacctgtcc tccattattc tgtgccaaag gagaccggtt agtgagttgg 96000
cctctatgtc ccatccaagt gttgccttta aagaattcac gagtgattcg gcgcatcgtt 96060
cagggtcaag gcttgagcta caccccccag ttcccgcgtc ggtaactagg gttaaaggtt 96120
tggcagattt tgccgcagag aaagaagttt cggtgttcat tgttaaagaa atagcttcca 96180
agaccccgga ttggaaacaa actccgtaaa ttttaacacc ggtaaagcag cgcctttaaa 96240
gtgaaggctt tgaaaagatg gttgtaaacc ggaaggcacg cttccaagtc tgcaaactac 96300
gccgaacgca agcctattta tatacaggtc attctgcagc tgaatgtatt tggtgcgaat 96360
cacgccgctg taaaaatccc tcaattgggc agctatttca caatatcctt taccagactt 96420
aaaaaaccca agtcctagcg ccggtctatt atttacccta atgtaggtta agcaaacata 96480
aatagacgag tgggcgggca aaggctcgtc gcgcaggctg gacgggcatg agtcatctaa 96540
atccacaaac atgtcactag gaagcgtaag gccaatatgt gtaacaacgg gctctctggc 96600
gactacgttg ccctttaacg cagacgtcac cttggtgaca aacgtactgt ggaccgtttg 96660
aaccaacggc ccgaccggcg caagaaactg atgaagcgag ccggtttcca acaattcttc 96720
aaaattgggt atggcgtcaa gtagaccgct ctcgtggccg taccaaacac acgctattct 96780
gttggtctgg ggggcagagt ccgcgtccat cctagacagt cgcgccagcg acgtaggcgt 96840
gaataacatg tcaatggagg acccagtgtc agtctgttta aaggaaaaca ggtaggtgcc 96900
ccgaggttcc tgtgaactca tggtctgaga ataaatcaaa aaatctccat acgtttgaca 96960
tgtaggcgag acagataaaa atccatcttt gatggcctcc accccagtgg tggtcgacac 97020
cacatattta gagagcagat cacgaacacc cttagaaaag tcgcgaccgc gagataacga 97080
aacccggtga ggaggcggcg gcagtagacg catcaaacta tcattcagct tgttcacgtt 97140
tgcgtccctg gtcttcatgg cgtcgtcaat tcctgccgca cgggctgata acggtgacga 97200

gaacaccggc gggctttaca aattaactga caacctccta acgtgcaccg gatcgctaca 97260
acagcttaaa ctcctgatgg agttccaact aaaaccacta ccaaccgcac accttttaag 97320
catgcccacc gtgacccggt ttttaaatac tgcattcaaa atagacaacc ccctggtttc 97380
atttattcaa aaacaccctg tgtttttttt aatgagagtc gccaggcttc cggagccagt 97440
cattacagac caccaaagcg cagaaacgtc aacaggcata ctatccgagg ttgtgaatgt 97500
tcttaataca gctattcgaa aaccacacga gtccccagcg gctaaagaca acgattatct 97560
cgacaaccgt gccatattgg ccatgattac agaatacatc catcacgtaa cttcacgtac 97620
gccctcgggg atcccaccga caccaccaat gggtatcagc catctaccgt gcgtagagca 97680
aattttacac gaaacccacc ggcaatactg gaacctaacc ctcccggagt cgctatttat 97740
cgacatcggc gaagtcgcgt ctccgcttca gacgtggcta atcctatcgt attgtaaaaa 97800
actgcaactg gcaccaccgc ccctattccc acctgtcgac gaactggccc gtcgtctggt 97860
taccggccac cacgaattgt ttgtccccct atcgacgtca ctggaaacgt acatcaccat 97920
gccagtatca aaacagcgag cgtttgagat atacagcgtc tttgccaaat caaaaaacat 97980
agttgacggc acaccaattt tagcgttcac cgacacagaa cttacaacgt tcacccccga 98040
actgttattc ctgtacgact tcgtaatcga gtccctgtgc aaaaaccagg catacgggtg 98100
ttcgcgcaac gcaattgagc attttattaa gaaaggtatc gattttatgg cggagttggg 98160
ggcgttcatt gaaaaaacgt gtggctatcg atcaaccgta agcctgtcca acgtgagagc 98220
cgttaaggcc aggctcgcga gctgtggcct atccaaagag gcgtgtgaag atttccgcgc 98280
aatgatactg atgacaccgc acgagacgac gccaaagtgg gaaaatttta cggacttttt 98340
agaaatggta aaccaattaa cactatacgg gttctatttt tacgagtgtc ttaaccagta 98400
cagccccact agcatatcgc tggccaaaat tcaaaatatc ttaaaccgag tcgatgccga 98460
acagagcgat cgcgccctgt ggcgcactcc attaatcggg tctttcccgt tcccctggaa 98520
attgaacaac gtgctggcgt tttttaagcc tagcaccccc gtggcaacgc tacaaaaaat 98580
ctacaaggca ataccgtcgt acctaatgag gtctttgttc gaaatagcgg ccaacaaatc 98640
gtggggtaac atcgcgctcg cagagagcgc cccgctaacc gacatacaga ccgccgaacc 98700
agaccagggt cccgtgtccg cccaggttat cgcaaaatat tgcagccgtc tccaaataag 98760
cgcgacagat tatgacgcag ccatcgtctc cagcccggga ttcgccgccg agttcattaa 98820
aacaaaactg tatccgatcc ttagcgaggt gctccgaaac acgtccaaaa aaaatcgatc 98880
gctattccaa attagatggc tcatagtctt cgcggcagag gacgccagag acctagcccc 98940
tatcagacgc tcgctggccc tggcgtattt tcaaataatg gacattttgg aagaaaagca 99000
ttctccggag tcgttttaca acctgctgga ctatcttcag gaaacattta ggtgcatacg 99060
acaggtgata ccggaagcca cgtgcccaca agaatttcta caatacttgt ttacttttca 99120
aaacattcca atagcagcga gcttcattca aacatccatg acctttgtag acgacctgaa 99180
aaacggcatc cccggtatac tggaccttgt ctccttaggt gccgcgtttt ataacatgaa 99240
actactgtac gattcaacgc tagacaccgt cgagattcca acggaagaag ggcaacccat 99300
cgtcgtgagc atgttcgtat tcaaatccac gattcgcgtc ctcgagaagc tcttacagga 99360
agccgttatc gcgttaactc aaacgtccga accgatgtac gcagcgcaca tccgtctgat 99420
gcaacacctc acgtacatgc aaaagatcgc cggacacgaa ataatgacca cacaacttcc 99480
atccgtgttt cacgaaatac acgagggata tttgcagtgt tttaagcgtt ttaaacgtct 99540
catgttacac gttacgggaa gctgctgcta ctcactgacg cgatactttg gattcctata 99600
tcaacccccc ctaatacccg ataccatcgt acaaaaaatt ttaaacttta acgacaaaac 99660
ggacacaacc gacgacatct taaagagcct gtcacagccc gtgagacaag gacctctatc 99720
ggctgaaaac gaaagtagca gtcgactctc aaaaaacaac gttgagctgc ttcaaaaact 99780
gtacgacgac tttcggaccg cctccacaaa caataacccc acctctatta aacttgaata 99840
ttcgggtaat tataacgaaa cacaagtgtc cgtagattgg agcacatata acctggtgac 99900
atacaccgca cccgacgata cgttaaaatt caccccggtt aacacggagg cactagatcg 99960
catgtttgca gaataaacga tggaattgcc gccaattttt tcgaaattta aaatagaggg 100020
cgtggcaacg acgcaccagg ccgactgcag gttcggacaa tacgccggct cgcagtgctt 100080
aagcaactgc gttatctacc tcgcacaaag ttatttcaat agagagtccc ccgtgacgga 100140
cactaacgac ctcgacgacg ttttacggca aggggcgacg ttggacttta tactgagacg 100200
gtccgggacg ctcggctata accaatatgc ccaactgcac cacataccca gctttatcaa 100260
gaccaacgag tggaccgcgg ccatctttca atctcaggag tactttggac taattggact 100320
tgacgcggcc atccgcgaac ctttcatcga atccctaaaa tcgatcctaa cgcgaaatta 100380
cgccggcacc gttcaatact ttctgtttat atgcggcgac aaagccgggg ccgtaattat 100440
caaaaacaaa acgttttacc tgttcgatcc ccattgcgta ccccacgtac caaatagtcc 100500
ggcacacgta atcagttcgt cagaccccac cgccatacta gagtacgtgt caccgccaga 100560
cagagaatac actggcagtt ttctatacat tatgcccagc gaatatgtca atccagagca 100620
ctacatcacc aaccactata gaactataac gttcgccaaa gtacacggcc cgcacataga 100680
tatatccacc ggcatagagc cgtgcaccat cgaagacatc ccaagcccgc cgcgatcgcc 100740
ggatgtgacg tcaaaatcat ccaacctcgc acgcgtgccg agaaccacca ccgacacgtc 100800
gagcgccaaa cccccaccgg cgacgctatc cggtttacgg ggcgcggaac caccgacaag 100860

ctacccagac ccggcaacca acgacgcgga cacaaaactc ttaactcccg ctccagcgca 100920
aaccgccgtc gaccaccccg aattccaaac aacacctgga gccacgctac ttctttcaga 100980
actgtcggca tccaggggtc ggaaacgcaa gctttccagc cttcagcgat attcggattc 101040
cgacgaagcg tcgtctgacg acgaaggggc cccacgtagg cgcgtacacg acgatgcgat 101100
atccgccgag gtgatctgga tggacgacga tatatctcct ctatattcac cgtcggcgac 101160
tccatcgttt gacgacgtgt tcgacagccc cccaatgagc ccggagttta catacgaaga 101220
cgcgacagag gacacggatg gcgcgtttct agaacagatc gctcgagacg cggaaacgcc 101280
gttctctgcg ttcgacgacc ttataacgga ccacgacttt tcttcccttg ataaaaaaat 101340
agaacagtta atcaagtacg aagcgccctc gcagcacctg ccaaacatct cggacaaaca 101400
aaacgggcga gccgtccgag aagcggcggc cctccaggcg atggacaaaa ttatgattaa 101460
tatcatactt gaacacggtc taattacaga cgcgcaggcc cggggaccgt ccgcgtgcaa 101520
aaacgttctt caatttttca tcctgtgggg agaaaaactc aacataccaa taagcgacgc 101580
caagcaggtc ctagaactcg atctgcaact gataccgtta catacggcta tcagcgaagg 101640
aaaattcaaa caggggggcgt tcaaaaaaca cctaacaact aaaatcaacc ggtgtctggc 101700
gtcaatgagg gccacgcacg cagatgcgca aaaaaaactg gcgtcggctt tcaacgtcga 101760
aggctcacag atttcgtcca gcgaagcgaa aatatcagtc cgggcgctga aggaacagat 101820
cgccaaccac ctgagtccag gctttttagc ggtctactcc gcggacgagg taaaacatct 101880
acgggataaa attcaggacc taaaaacagg catcgagcag cgcaacaaag aaatccaaca 101940
ggaagaactg tttttgatg ccatgctcac agccctggac acgttccaac cccctccgaa 102000
aacggcattt ccaatggaga tctttccgca ccgtaaaacc gaagttatgc tcgaccacct 102060
ggcgtccata accaccaggt taaccgagga cgccaccgaa gccctcaaca attacctgga 102120
gaccccgccc gaccaaggaa cgcacattac caacattcca aacttttcat ccatcgtggc 102180
aaatatcata tccacgttaa aaatcctaac gtacgcagaa aacgacatgc aattaaacgt 102240
aacgcccatg gcaacgtaca ggcgtcagct gttgtacctc ggaggcgagt tagcaaccat 102300
ctttaattta gagtggccat acgaaaccgt gccaccggtt caagaactgc ccctcgtggc 102360
gcgggcgaaa gcaaaaatgg aatcggtaac aaaaatggaa aagaaccaac aggctctcga 102420
ccaaatactg ggagacgccg aaacgttact tgacacaata accgcaacat ccggagatga 102480
gaacccggtc cgcgccatgt ccataccgat actggagacc tacattacaa acgcaggcgc 102540
cctgataggc agttctcgaa accagcggtt cgaaaaactc aaggccgcca tccacgacct 102600
ggcatcatcg gagtcgttca taataatgct gctaaacaac acgcggctcg ataacatatc 102660
agacaatctg gccaagatcg acggcatcct gaccaacaac acacgttttc tttcaaacgc 102720
cactgttagc aaaacgctcc aaacgctggg aggcagccta atacgcgaat gcgtagaagc 102780
gctaaataaa aggagcccct cttccctcaa caacgcgcgt ctcctcgcgg ttcaaaccat 102840
actggggcac gcgtccgttc cagatcacga gacgctgacg cgaatcgttt ccggcgtcgc 102900
cagcgcacaa aaggaatccg ctggcgatga tccagatagg tggacgcgag taaccggtca 102960
cctaaacgag ctgaagctcg taactaccca atcgcgtgtc gacaaagcca ccaggcgcaa 103020
actgttaatg ataataaccc gtgacctcaa ggaggcggag gtgtctcagg aaacggtcct 103080
ggaaacacgg tggcaagaaa acgtgctaaa gtttcaaccg tcgacgtcca aagaaatcga 103140
agacttttta cagtcggcac cgtcagcaaa ggcccgaaaa ttcgcagaaa aacacctacg 103200
gacgctgatc acccaattca acggccacga gcgaccgccg tccgaggcca ccgccgttcc 103260
catggactac acgccgacgc ccatacccac gccacaggcc gtttctacgg ctaccgcgga 103320
aaagggaaag gccgcatgga ataaaattca acaggccttt caggatttca actttcacct 103380
catcgacgct tcggattggc aagagatggc atcagaatac tccagacacg gctcgtccct 103440
tcctggtacg gttggaccaa agctggtgcg cttcatggag agcatctcaa acaccctgga 103500
cgacatcctc acgcagaagc tggcatctct gcttccaaac gggcccgcgt tcagaccccc 103560
agcgtttgac tggatcgcgc cctatcaaac acgcgtaaac gcgtttctaa aaaccatagg 103620
cctgcccatg gtgcgcaacc tggcggacaa gatccatcac caatgccaaa ctgtcagtca 103680
cgcggtgcaa tccgcagacc ttcaacaggc cacggtggga acaagtttag aacgacccgc 103740
ggccgaatac tgtcgaatac tctctgacat gcaagtcgcg ttcaacgacc acggaatcgc 103800
cgtaagatcg gaggccgcgg cgtacacgga cgcaatcaac tcgccggcca acgtcgtgac 103860
tcccccgaaa cccaacctag aagcccccaa gaagctaata acggcaactg acgccctaac 103920
cgtcgaggac tttccagatt tcctaaaaac gtcaatcctt caacaggagc agcgactcat 103980
tgcgctccag agagcggaat ttcagcaact agaggccagc atctcggcgg ccgaacggct 104040
ccgccaatcc acccgtgacg agatcgcagg caagatggca accgctatca cgcaactctt 104100
accccgcgcc cccgtcgcaa tatcatcgag accgttgaac ttatcaaaac ctatagactt 104160
tttgagttca acggtatacg acaaaatcct ggacaaggag ccttacgaga cagccatagc 104220
gggattcgcg tggctggaaa tcgcgacaaa atccgtaatg gtctacagtc aacaaaacga 104280
aacgcaacag ttaaacgtac tgctgagcga ggtagaaaaa cagagcaccg tcgcgcagcg 104340
tctacacgat ttggaactgt cggcgaaaaa cacggacgac gtaaaggtgc tgaagcaggc 104400
gctagacgaa ctcgcgcccc tcagggtaaa gggcggaaaa accaccgtag acgcgtggaa 104460
acaaaaactg gaaagcatag aatccctgct tcgcgccacg aggacggcag gcgaaatatc 104520

atcggagctt gaacgcatcg gcacacaggc ggttggcacc atcaccgtcc gcgatttagg 104580
aacgctctcc gatcaatgcc gggaagccgc aaatttcctc agacaggcca gtctacccga 104640
aggcttctcg gacataggca caaaactcag cgagcttcag gcgtacatta agtacaaaaa 104700
acagtttctg gagcattttg aaacaaccca gcctaacgtc tttcaacgct tcccgctatc 104760
ccaaaacata accgaaaacg ttccggcgcg cccggcgatg gactcggtgg ccagactgac 104820
caatcacctt cacgtgcgcg gcagcgcgcc ccactttaca acgtggatag aaacgctacc 104880
gaccgtcgat ccggaaaaac caactcacgt cccggcgcac ggaggagccc ctctgcaccg 104940
ccagatcacg tactcaaacg tcctagaggc gttgttttca ttatgttcca ccacgctaac 105000
cccggttccg acggcccccg gtctggaaat cgcaaccagg gcacgccgcg gggcagaggc 105060
cgcaacgtgg atggacagac agtggcccga catcgctcag acgctccaag acgttctcga 105120
cacgtacgaa cacaccaccg cccacgcaaa ccgggacgcg gcatttaaca cattcttggc 105180
gatgtgcgtt tttacgcaaa tcatcagggg cgctagcaga gccgtgacgc tcccgaagtt 105240
accgagcacc gccgtcgatt ttccagaaga gatcgttcta acacccaggg aatgcacaac 105300
actggtcacc gccatgtggc ccaccctggc ggccgcaatc ttacgattaa aatcctactc 105360
ggaagcccta ggactaatga gtcgtttcct cccgctaatg ttccaggcgc tgccgcacct 105420
aacgctagag gcccaggtca aaaacggccc acataacacc ccgcctcagt tgagatgctt 105480
tgccaaaaca gaggcaattc cgtatttccc ggcgcaatgg cagtcagcga acctagagca 105540
gagcctgtgg ggacagacgg actttttgca aatctgcgat aacaatcaac gcaaggccag 105600
ggtggcggcc gtcacctggg cgctcacgac gatagacggc gtggtttgg accaactgtg 105660
gtccacattt aaacccatga cagccgcgtc agacgacacg tacgtcgacc tagtcgagac 105720
cctacacctg accacctttg gcccgcgcgg tccaacgccg aggcgagaaa cgaccaccga 105780
gcacccgccg tacgagtacg gacagcccac gggctactgc atctcgggtc aatcgacgac 105840
gccggtccag gcttcaaaca caccggtatc cgctttcgag gcggtgctcg gagcaatggt 105900
gtttcacgta ccgatcagaa tatttttggc ggccacgccc aagcgccttg gccaggcgcg 105960
cggcggcatg gggctcctca cacccatcct ggaatgcgtc cccgacgtcg agcccttcaa 106020
aagcctgtat aacgcacccc gcaaacccgt gcccattgaa acgctacccg catccctcca 106080
cccgcacgac gagcgacagg tctttctgag acaggcacag tggctatcct accgattcac 106140
accacacgaa gccgcccggt cgtcgactcc gccgcttctg gtggtcatag accctgaaaa 106200
cctcgtaacg gcaacgtact ccagtggcgg gcctgcaaat ttcgagagca ggccgtttta 106260
cgtgatgccc ggaccatacc ccccagactg gccaaaaacg ctgtcggtaa catcaaacac 106320
gtccgtgacg cacctcagcc acgacgagat atgtaacctc tttactacgc tatcccgaga 106380
acacgggacc gtgcaaggca gggatatctt cgcagcggct ccgacaaacg tcacaccgga 106440
acaaaccgcc aatcctccgg catgggaaac ggataaccga ttaataacgc aaacagaaac 106500
cgccaaaaaa cctcatataa ttcctgcgtc tcctaaagcg cggacagatc caccggtgga 106560
aaccacgacc caccattcac aagggcaagc gtcgcaacac gcaaacagca acgtaaacca 106620
gcccggtcaa attacttcac acgcgtcacg taacacaccg tcaaccgcac ctcaggcctc 106680
atcttcaccg gaaaaattca acacgcaaac ggtgcctcga ctaatatctc aaacgtcgga 106740
aacggcccat ataaaccagc cagcctccgg ccaggtcacc gaaccaaagg gaatctttgg 106800
gacgtataaa ccccgagtgc tcaccgaacc cgccaaaccc gcaaacgccg gcgtagcctc 106860
tcgccaacca gaggcaacca ccacggtccc caagttaccg attaatccac ccaccgctag 106920
ggtctttata gggaccgcgt ccaaactctc gccagccgtc gaagagagcc acggcgccac 106980
acccgacgca catcagtcga agatagatcg ggaaaaatac gccgagagtc ggcctcgccg 107040
caccccacac ctcgaagagg ggccacggga gcctcacgtc aacactccaa ccagcgcaca 107100
cataaacgtc ccctctagcc aaggtcaaaa aacagtacac gggcgcgaaa atcccggcct 107160
tcaaacagca actcccagcg ccccccaacc aaccgcatca aacccgcgca ttcaatacac 107220
gctccccaga acggacggcc ggttgcttca cgacgaatcg gaggtggaat cgaccccaac 107280
cgaggaggta aaacgatcgc caaaaacaca agatgtgtct cacgggcccg aaccggacga 107340
ctccaggtgg accgccccgc tcggtccaac catagagatt catcgactgg aacaccccca 107400
aattctcaaa aatataacat cactcaccgt ccccactccc agagtcaccc caatccctcc 107460
cactaacatc tggatacccc tatcccacgt caacatccaa cacgaagaaa tcacacgagc 107520
caagaatgtg ttaatgcgat ttattcaaaa cgtacgaaga aaacttcaag cgtcgtctga 107580
cgctctatcc gaggctattg ccagaataaa gtttttatat ctgtaacgcg cccatctcac 107640
ttgctttttt tattttgaga cgagcgtctt gtgtccagag tagttgcgtc gctaggtgat 107700
aacgaagtgg accctatgcc agaagacgcc acgctcccgg gtccgccgcc accggcgcg 107760
gggcccatcg agcctatcaa tgaatggggc ccgctggaga tcgtagtaaa gctatttgac 107820
ccgagggtgg aggccaccgg ggcaacgctc ggggcggccg agccggaccc ggacaaaacg 107880
cccgagatac tagaactagc gtcctttttg ccgcgaaggc cccggaggtg gtctttgcgc 107940
agaattccat tcttcttttg catatacatg tcgtaatgat gtttggccgt taaaaacacc 108000
agataattac gtttcgcgat ggcatactgg gcgggagaca tgtcaccctg gggaaggttg 108060
ttcatctcgg caaccagcgg gtgatttggg taatcgtgct caaggcgtcc ctgaacgatt 108120
ggctccttaa cccgcaagga cgacatctct tttaacttct attacacttt ccacaggaca 108180

gggacgatat agacgaggtc aaataaaaca gctcggccac acgcaaatgc tttaataggc 108240
cggtcgcgca gtccggcgac gccaaacagg gcacgacgac gctaaccagg gagttcgcgt 108300
ctcgtatggc atgtgccgcg ttgttggcga gcacgcaccg taaataggga tctccaacac 108360
acgtgatctc gaatagagat ataacccgca tgtgcgatcc gccacaataa gagcaatata 108420
cgcgcccggt ggtagcacag atcgaaagct gcttctcttt ttggtcgcga ctgaaaaaca 108480
cgttggtggg tgggaaattt acggtttcaa atttaccccg tccgaaattc aaacagtaac 108540
cgcactcgag gcacaccacc accctcggag ctggcacggt cttctccagt acgctcctgg 108600
ccaccacctg ggaccaaaca ggtagagaga tacacggaaa cagtacgtta tacgccaata 108660
ctttttgacc caggtcgcgg gatatctccg tctcggtcga ctcccctatg ggcaacacaa 108720
cacgggacat gctcagcagg gccctaaacg tcaggctcct cagaagggcg ttaaacgggt 108780
tgccgcacgg gacggtcggc gccagttctc gcagcgaggc cagaagtccc gcgtccgaag 108840
ggcccgggac actctcattc aggttagctc ccagacgtct ggaaatggac ggaacgttca 108900
actgcatcga gacacaaccg cccccgttcc atttcttccg caaaacgggc agatccaacg 108960
cgtgctgtgg caaacaggta accaggggaa accgctggcg acagttaagg gttttgcaca 109020
cgagacaaca cgccctctga aacgacacaa caagatacct ggaccacgcg ctcctgggaa 109080
cggccgttaa acttaaacct tcgtcgctgg gacagcccac gccggtgcag atacacctca 109140
gcacccacgc gtacgcctct aacaacgacc ggccgatatc gtgcagccgc gatctcacgt 109200
cgccgttttt aggtgggtta tccggcctcc agccggtagc aatctcgttc agggcggtct 109260
gaaaggatgg ggcagaaatt aacgtgcaag cccattttgg gggtcgtccg tcccaggcac 109320
cgagcccgta cgtcacaaaa cacacgtagc attcggggca tagtccgatt gagcgtatag 109380
aggccgcgag atctaagccc agccgagaac catccagcca acgatgggca ggataagaac 109440
cgcgcccatg acaggcagct tcgtcttcag ccaggttagg caagcggccc gaggccatcc 109500
cccaattgta ccgattgaat tggttaattg gtcatcgccg cacgctctcc gcgcccaata 109560
tccttcaact ccgaccccga agggcgggcc acggagcccg ttcgcctcaa cgcaccgagg 109620
ccgtcggagt gacccgtcag ggcaaaaacc gttctaagaa gggtttttaa ccgtttagcg 109680
ctctttggag tcacgaccaa aaactgtaaa acctgtcggt gctccgtaaa gtaggtgcgg 109740
catatgacca tggagctgta aacgtttagg tctccggaga aaaccagacg tgccttaaat 109800
ttcataaaat cgtcctggcc cagggacacg gacgagttcc tctcaagata cacgtccgaa 109860
tttataggca ggttttttcc aaactgggca tcggcgtcac gtggcttaca caaaaaacat 109920
ttcagcgtgg tggccaaacc gttgttgata attacaaaac acggggcaaa cgggtaggcc 109980
agtctctcta gtttgtggag ccaaaactta tacacaaact cgagatgata gacgcagccg 110040
tgctgcaggc gcacggtgca cacggggacc gccccgcctt tagcgtatac gggagccccg 110100
tcctgacacc tctccaagtc cagggagatt ccagagggtc ccaggtaaga gacaactaaa 110160
tcgcacagct cgtcaactaa acgttttccg gaactcatcg ttataaagat cctttaggtg 110220
ctgtgcgtgg ctcccgtaaa aaccgcgtcc gtgctaacga ttttgtgaat gacctgtttt 110280
acggcgttta ccttggcgtc caggaccatg cagtgctcac agtgagctga ccgcgtctga 110340
gcacgatgac agaggaaagt ttttaaatac tgacagtagt taatggcgtt gagcctggaa 110400
tatatggtgg gaaacataat tttcatgtca tcgggcagca gggactcgaa cgccaataaa 110460
tcgtcaccga acatcacgtg agacagaggt aaaagatgct caccgccggt accgcgtaac 110520
acgcgaccag ccaccccttc aaatatttta gccttaaaaa gcgggccccc taaagtcgtc 110580
caactcagct taaaaactcc tacccattat gttcgtcccg tggcaactgg agaccctcat 110640
gagacactgg ccctctctgc ggggactcgt agaacaatcc ttcctccccg gtaccccgga 110700
cggagctttt aacagcccgg tattaatcca cactcaggac tctctacaac ccgcctcatc 110760
gtgcagggtg tgtagcctcc tgttcactct ggtccggaca ttcccacccc ccgactcttt 110820
cttcgaagac tacggctggt tgtgcctcac ctgcctatac gccccccgat catggacggc 110880
taccctcatg gtggctgccg accttttgga actaacgcac gtgtacttcc cgcaatgcgt 110940
gaaagatggg ccagtataca ccgcccaaag catcctcgga atcgacgtcc agctgcactt 111000
cttcgcaacc cgctgcttcc gacccatcga cagagaacaa atactccaca catctcattt 111060
aaatttttta caaaccgagt ttattagggg catgttagaa ggcacgattc cgggatcgtt 111120
ctgttttaaa acgtcctggc cgcgcacaga aaaggacgac caacaaccta ccgttgcgtg 111180
ttgttccgtt ggccgcggaa gtcacaccaa ccgggataac cgcctacccg aggacctgga 111240
agaggcgttc aactccacga acgccgagga aaagcccagc ctcctcggcg tcttttcggc 111300
aacgtgggca gaatcccagc ttcttggctc cgacacacaa caggcagata cccatttaca 111360
accctccgcc ttcccaaccc cagaagatgc tgaccaatca cagggcccct gcctgatgca 111420
cccaacgctc aacctaaaaa caaaaaacca caccgcatcc atatgcgttc tatgcgagtg 111480
tctggccgca caccccggacg ccggtccggt tctgaaagat ctgcgtcgcg acattctgga 111540
aaacatggaa aacaacgtta agctcgtcaa tcgcatatcg tacatcctaa acgatccgga 111600
ctcactgtca cacgtgcgcg acgaacatct gcgcggccta attaaacggt gctcggcaca 111660
agaaatccac aagcattttt tttgcgaccc ggtgtgcgtc ctgaacacgt actcgcactg 111720
tcccgcggtt ttatttaaat gcccacctcc cgaaaagtat aagaagctca aagctcgtct 111780
ggcaaccgga gagttcctag actgcaacag aatatttgac tgcgagacct tacagaccct 111840

ggccgtcctc tttaaggggt ctcaactggc caaaatcggc aaaaccacgt cgctcgagat 111900
aatccgtgaa ctcggatttc aactgcgtcg acacaacatt caaatcaccc acccgtttca 111960
aacctccaac ctatacattt aatcttcaga agcgcaccag acaatgccaa aacagcccag 112020
aagtcgattg gcgtctcgag cgccgtacgc acctagcgtc aggcgaccgg acgggcccca 112080
gtccacgcga ccggcatcca ggcacggcag ctgcaaaagc gaaatcatgc agtggaaaaa 112140
gttagtttca gacacgcagt ttttttctgc cctaacgcgc cgccacgagc tgggggtgga 112200
ctttttaaga gaaatgggga ccccgatatg cacctcaaag tccgttatgt tgccgttaaa 112260
cctaaaaacc atcgccccgg gtcggtgcgt ctctctctca tcattcggac actcgtcaaa 112320
catggggttc aactgttcgt cgtgcacgcc aactgacagg tcagcggtgt ctctggacgc 112380
aaacgcgctc ggcgaagatt ccgccaggaa aaacagcgag ctgtgttcag tggcgttaac 112440
cttttaccac cacgccgaaa aggtcgtgca gcacaagggc ttttacctgt ctctgctcag 112500
ccactccatg gaagtcgtta ggaaaagctt cacgcaaccc gggttgctct acgcccacct 112560
agtcctaaaa acctttggcc acgatccttt acctattttt acagtcgatg ccgatgagag 112620
actcgcactc tgggcggtgt tccacactag agacctacac ctgggggaaa ccagtctgcg 112680
actcattatg gacaaccttc caaattatga cataacggtg gactgcatca agcaaacgta 112740
cataatgaag tttacaccct cgcgaccgga caacgcaacc gtgacggttc ctgtcaacag 112800
catttgcgag gccgtggcca ccctagactg caccgacgag tttcgagaag aaattcaaag 112860
gggcacggcc atcataaact cccaggggct attgtaacct ttccctagac ggaaacagat 112920
gtaattcacc tattcaaatg ttaagttttt gttttgaaca tatcaccaat aaaaacaact 112980
tatgtttaca caaattgaag agcgtttcaa ttttaccata aacatagcaa aaaccacggt 113040
aactaaaact caatagcgat attcggatgg gatccctaga gagcaatcga cgtgcgccgg 113100
tattaattga cgggaacgca cgttaactgc tctccacgaa tcgcaaactc cgcgttttta 113160
ggaccattcg acgccgttaa taattggcag taaactgcga aagggcgtac atcgcagtag 113220
tgattttacg atgtgcacac ctttccgcgg ttcaccgcca acaacaaaag cgggtcacca 113280
ggcacctaaa tcgcccgggc cttttatgcc aaagtacaaa agggaccggc gttctacgtt 113340
cacgaacaac actgttgatc gcaaatacaa tagtgttgtt ggtgaccgca aactgccaaa 113400
tcctaaaatg catggacata gtcctcaaaa tctagaacga acttccaaaa acatgcggca 113460
ataaacaacc ccttgcagtt aatgcacatc gtaacgagac ggaatccaaa gaacacgccc 113520
gatcttaaga ccacgggcac gtgtctttgg gttccggggg cgttacgatg gaagttcatc 113580
accaagtatc ccatcaacct gttacaaaag cgtaaagata atcggtttac aataaacaat 113640
aaaaaatagt ggacgaagtt acattccgac atacaagcgg acccaaaaaa acaccccaac 113700
gcaaaacaac tggagactgt gtctttgggt cccttccacg tcgcaatgag atttctggtg 113760
accgtgtgct gcttaacgcg ttacagttgt gttttatggc tacgtaaaca caccaatcta 113820
gaatgctaac acgtgcaccg cgcttaggcg cgtttgtgca actaactgct gtcaggttac 113880
tactcttatt tttaacaagt agcataccaa ctgcaattaa ttgcactaaa cccagtttag 113940
tctttttagt cgagtcggcc catagtcttc atggagaacc atcattcatc tccgtcaacg 114000
gcgtcggtgt taccgttgtt attgcgggca aaacaggctc gggggaaaaa agagcaacat 114060
atcgtcagtt taaaacaact caagttctaa aatctcaccc actcgtaaac aaaaaacatc 114120
gcagttaagg tgtatgtacc gaacataaca acacaagttt tttaaaacac agctgcggta 114180
agtaaccccc attgccacgc gtgtcgcgtg ctaagtgttt ttaaaattac attgtgcgtt 114240
ttacacaccg agcagtaatc tcaggagggc ggttaacgag cgatatacat attccctaaa 114300
cacgggaacg cgcgctgacc gccctcccca aatcacaaca cgggactaca aagcctagtg 114360
ttaatataat caaattaaaa aaccacagaa acctttagtc gtgcgcaaac actagcaaag 114420
gtacctagag ctttccccta tacttcaaaa aacagcggtg ggttatttga ccacacgtta 114480
agtaaacacc cgtaagaatt attcccgttt atcaaaatgg aaaaattaag gctttgcgtt 114540
aaaatctgct aacgcaaagg gcacttaatt tttccagttt ggactcggaa ctttaccgct 114600
aacgttaaaa tttaataatg caaaaggcac ataatttttt ttgtgtaacc tcaaaactct 114660
acagctaaag ttaataacgg cgactttggg cagcgaatca gtgtctgtcg caaaccctgt 114720
taaatttaaa cacaacaggg agcgcgacta aacacggact aactgctcat cgggccctgt 114780
tgaagaagtt ggtgttgcaa tgcattaggg aatctaaaaa cgaaatcctg ctcggtaaca 114840
cggaaattgt ctttagttcc ctacagcatc acaacaattg taaaccataa acgtacgcgc 114900
gtggggttgg tgtggtgtga aatgcttggt gcaactgtta catgggcgga ttgtaaatgt 114960
ggtgcttggt gcaacggtgg tggggtgcaa gtcccccggt ggggtgcaag tcccccggtg 115020
gggtgcaggt cccccggtgg ggtgcaggtc ccccggtggg gtgcaagtcc cccggtgggg 115080
tgcaagtccc ccggtggggt gcaagtcccc cggtggggtg caagtccccc ggtggggtgc 115140
aagtcccccg gtggggtgca agtcccccgg tggggtgcaa gtcccccggt ggggtgcaag 115200
tcccccggtg gggtgcacgt cccccggtgg ggtgcaagtc ccccggtggg gtgcaagtcc 115260
cccggtgggg tgcaagtccc ccggtggggt gcaagtcccc cggtggggtt caagtccccc 115320
ggtggggtgc aagtcccccg gtggggtgca agtcccccgg tggggtgcag gtcccccggt 115380
gggagcggct cggctccggg gtggctccgg gtgggggcgg ctcggctccg gggtggctcc 115440
gggtgggggc ggctcggctc cggggtggct ccgggtgggg gcggctcggc tccggggtgg 115500

ctccgggtgg gggcggctcg gctccggggt ggctccgggt gggggcggct cggctccggg 115560
gtggctccgg gtgggggcgg ctcggctccg ggtggctcc gggtgggggc ggctcggctc 115620
cggggtggct ccgggtgggg gcggctcggc tccggggtgg ctccgggtgg gggcggctcg 115680
gctccggggt ggctccgggt gggggcggct cggctccggg gtggctccgg gtgggggcgg 115740
ctcggctccg gggtggctcc gggtgggggc ggctcggctc cggggtggct ccgggtgggg 115800
gcggctcggc tccggggtgg ctccgggtgg gggcggctcg gctccggggt ggctccgggt 115860
gggggcggct cggctccggg gtggctccgg gtgggggcgg ctcggctccg gggtggctcc 115920
gggtgggggc ggctcggctc cggggtggct ccgggtgggg gcggctcggc tccggggtgg 115980
ctccgggtgg gggcggctcg gctccggggt ggctccgggt gggggcggct cggctccggg 116040
gtggctccgg gtgggggcgg ctcggctccg gggtggctcc gggtgggggc ggctcggctc 116100
cggggtggct ccgggtgggg gcggctcggc tccggggtgg ctccgggtgg gggcggctcg 116160
gctccggggt ggctccgggt gggggcggct cggctccggg gtggctccgg gtgggggcgg 116220
ctcggctccg gggtggctcc gggtgggggc ggctcggctc cggggtggct ccgggtgggg 116280
gcggctcggc tccggggtgg ctccgggtgg gggcggctcg gctccggggt ggctccgggt 116340
gggggcggct cggctccggg gtggctccgg gtgggggcgg ctcggctccg gggtggctcc 116400
gggtgggggc ggcctaaaat ccttaccggt aaatttagca gtaaatccaa cgcagtaaat 116460
ccgcaagcta gccgcacaga ggtgcgactg cctgccaagg ctcctggcgc ctcttttata 116520
gcgctaaatg ccctccccaa atggttacta tggtttagtt atgcaaccat aataccaata 116580
aatgaatcac caagaatacg gccaacgcgc cagcgtccaa ccgccctcca accgtgcagc 116640
gccgtatatt cgaattgcct tccgcacgaa tacgggcgc agcatggtcg cgggaggctg 116700
gcattgcgca atattcccag cattccaacg ggcacgcaag ggtccaggct accggggatt 116760
ccccaaacat catccagcgc atagaggttg tggctgagct gcctatctgg ggattcccca 116820
aacccgagac ccactgagat gctattcagt gggactgcag gctgggtcct cccaggacag 116880
gtgcctgcgt gtgtggtaag ggagttcccc taattaaaat attatattaa taaataaacc 116940
aggctaagct gtaattaatg aggcacaggc tgccgaccta ggtacctggg gaacccccta 117000
acgggatgta attaattcaa acctatataa attccaccct gttaggggga tccccacgtg 117060
tacctgaatt aatgcgggat acaggctgta ggtttggtaa tgggtaccag gctggagtga 117120
ctcataccat gggattaaat accagggata caagctgtag gttcggtaac gggtaccagg 117180
ctggagtgac tcataccatg ggattaaata ccagggatac aggctgtagg ttcggtaatg 117240
ggtaccaggc tggagtgact cataccatgg gattaaatac cagggataca ggctgtaggt 117300
tcggtaatgg gtaccaggct ggagtgactc ataccatggg attaaatacc agggatacag 117360
gctgtaggtt cggtaattta aacctatata aattttaccc tgttggggga atccccgttt 117420
gtacctgaat taatatagga atacaggctg tagaatcggt aatggaaacc aggctgtagg 117480
actaggctgg agtgactcat accatggaat tcaattaaca aggaaaatat aatagaatat 117540
atatatatat agggaacttg taaacaaaac ccaactcgcg gattggctgc ctgtttgggc 117600
caaccagcag cgagaattcc ctgattgaca ggcggactgg ccaatggttt gcgagcattt 117660
ttgattgaca gaacggccgg ccaatcgcaa ccgagaactc ggcagcgaag caaaagacag 117720
acggccgcgg cgaccaatgg ccgccgcggg ttagtttgat tgacggcttg gccggccaat 117780
gggaaacgtc cctgccggcg gcccctaatc cccctggctt aagggattac ggccgtaaac 117840
atcgctggcg cggtgccacc gccgccgaac ccccgcgccc ggttcggcgt gacctcgccg 117900
cgaccccgcc gccctcatgc gcaccatggg tgcgatgggc accggcggca acattattcg 117960
ctgaccgccg gacgaccccg gcgccaaaga gcggccgatg ggtacgcgcg gcgaccatcg 118020
aaacattttt caataacccc agtctgaccc cgcggcccaa atctggggcc catggtggcg 118080
cgcggccccc accgtcccca tttcccacgg tttgctcatg ggcaccgacg gtgcccgtgg 118140
tgatatgttt cagtgaccct cggatgaccc gggcgacgtt ttgcgcacca tggtcccgaa 118200
ttcccccatg agtacatgtt tcggtgaccc gccggcgacc ccgggcggca aaagccgccc 118260
catgggcacc ggacggcacc atgtgcacat ctttcagtga cccgcgggcg accccggcgg 118320
cggcaccgcg ggcgctgggg gtgatatgtt ctgatgaccg ccgccttacg tttgcagcgg 118380
tgtcagcgga acatattctc cccggctccc gagcccgggg cccgcggacc cgagttgggt 118440
tattaaagtt acgttaggga ttaattttaa tccactaccc cccttaatgt gcggctcggc 118500
ctgcggcgac gctgacgccg gcgggttcga cccattgacc gcatcctcgg tggccgggca 118560
gaggccggcc ggccaaaggt gcgatcgcgg ggtctgaccc acccgtggga gaaacccctc 118620
cgttccgttg tcttctgcct gcggtccctg gccccctggc ccaccgggga tgtcttgtga 118680
gcactttccc cggtgagaat agacaggaga gtgcccgcaa gacatcgacg gggttgccgg 118740
tcggtgaacg ggaggcgcgc cgcttgcgtg cgtgaggtgt cgccggtggg ccgtatgacc 118800
cccgcaacgc gaggggccgc tggcgcggaa atttccccag aaccggccgc gccgcttggc 118860
ggcgcggctt cctcccccaa cgacccctaa tttggttttt aagttgttat ttaaagctcc 118920
gctgtgtttg tggcacgcgc ttaagcgttg ccacctgttg cggagatccc ccgtgcgcat 118980
cgccgccctc ctcgcctcgg tgcacgctgc cgagcgtgcc acctgtttgg ttgtgtcatc 119040
ccatgtgcgc atccccccat cttatttccc cgaccaggca acataaacgt cacgtcgttt 119100
gtaagaaata attgcttttt ttatttttcgt taaccccgca accgcgatac agtctctcgc 119160

ctttcacccg ttgttttgtt accagcccct cttggcggcc ggcaccggcg ttaaccgggt 119220
gcgttggcgg ccacgacctt ggccacgtcg cgcctggaga cggcctgcag caggtcactt 119280
aaaaccatgt agtttgaggg gctgaccagg gccgcctttt ccatttgcga gagccaccga 119340
aaaaaggtgg gtgactggtt gcgcctgcga ccaagctggt cgcccgttag aaaaactaaa 119400
tttttcacgt ccctttctga gagctggcgg tcaatggaga gcatcaggtg tttgtatgga 119460
ctaaagtaac tggggctgga ccggagcctc tcgataataa atcccacgtc tagtaaaaaa 119520
aaggtggaga ccaggtccaa tcgcccaacg agaaacacgg cctccagcaa caccgggagg 119580
ggaataccgg gttctccagt tgaggggcac agcccgtttg caaacccctc gggcgtatcg 119640
tcggacgccg gccggtcaaa aagccacagc acggcctctc tgtcgtctgc ctctaggtgg 119700
cgaccgaagt ccactaaccg cttatgcggg aacatgggtc cggtgcggcg ctcgccgaca 119760
ccctcttcta cctacactag agcgtttgat aaatcacgac gaaccgcggg cgttacttaa 119820
aatgtggtct tcgtcaaatt cccgcaaggt ggtggcaacc gactcggcgg cggccgcgag 119880
ggttgcggcc gaaacgccga cggcctccgc caatctggcc ggaggcggcg cagcggcacc 119940
gaggctaaac agcgcgcagg cggccgccac cagggacggg ggcagaccgc cggtggccgg 120000
gttcacgatg gctttgtgaa cagactccac cacctggctg tgaagggcaa agagctgctc 120060
ttttgtaaag ccgcttttaa acagggtggg cccgatggcg tccgttggta agacggcctc 120120
tagtttccac cttaaagcct taaggattga tttttcctgt ttcagtaaat ccgccacgga 120180
aaaatcctcc gccgcgcaaa agcataaaaa ggcggccttg aaaggattca ggtctctaat 120240
cttgccggct aaaaatagac aggcggcccc cagcctctga aaccgccgac gggggatgct 120300
gcgacactct agatagcgat ccagaatact aacggccagt gggaaaacgc tagcatccgc 120360
ctggtgggcc ctggcaacag accgcatcca cgtccccaag ataactcgca tgcccaccgt 120420
cacctccgtt tgtatcgtcc catagcaggc cgtggatgtt acaaaggacg cttcgtgggc 120480
caggaggtta gagagggccc ggtcctggta caggacgggg tcaatggtcc cggtgggcac 120540
tgggccaaca gaagccattg tcaaggcaag gtaaagaaag cgatataagt cctacctgcg 120600
gctcacggag gcttaatacg aacgtccggc ttaacacaca atcttctttg cgctctatgt 102660
ccgctccaca ctggtcccat gtcttctgca gttcagccct gcctgttccc ggaaagcggc 120720
ggactagatg caacgcggaa gcggctgttc atttaaggtg agtcacaggc ggggcaattt 120780
gccaaccagt gacgcgtttc ctgcagtctg ggttttcacc tggcaagcgc cactggttct 120840
tggcagccgg ccggtgggga ttttttttagt gctgaattgg cagtcctctg tccatgcgca 120900
ctatgctggc ctggaagtgg ggggcggatg ggtgatgtcg ttcccaccgg cgacgggcct 120960
ttttgacggc cttggcggtt ccgcggcttt gacagaacac cgtaatgcaa aaagccggga 121020
tgacgggact gccaggtgcg tccggcaaaa cgcggacgct ggatgctcta atgccgggaa 121080
actgtaggca ccttgccagg cgatacagtg aataggggtc tgagttacaa aagatgactc 121140
cgtgacacca cgggcccggc cgaggggggt cgatcctggg tcccaatacc tgacgataaa 121200
atcttctctg tgggcggtat tttctgggta cctcttcctc tgatgaggga tacggtatgg 121260
gaatggggta tcgcggcggg cgcgggggc gcggggggcg cgccggttgc tgctgccgct 121320
cctccccctc ttgctcctgt tgtctttcag gagcttcgtc ctcgtgggag ttgtgactcg 121380
tgcatctgag acgtaaacaa ggaactcctc cggcacgcgt gggtggaccc caacccccta 121440
cggtgtacga gtgggatgta tggccaggaa acgggggtcc tggccgagac ggcccggtgg 121500
ggtggtctgt gtggctcacc gcctgttgcg tgttagggga tggggcctgt tgaggcgttg 121560
gcgattgggg tggtcccggt gagtcgggcg gggattgcgg aggggggcgag gacgtcggcg 121620
ttggcgattg tggcggtcca ggtgagtcgg gagggggcaa ggacgtcgga ggctcaggcg 121680
gttcaggtga gtcgtggtgc gatggttgca aaaacgacga aggctgctga ttgggcgatg 121740
gattatgcga cgacggcggt gcaggtgagt gggggccggg gtgtgacggg ggtgatggtg 121800
ggggcgaatg cgatgcgggg gatgccgggg acgaccctga cgtgttttga ggtgacaggc 121860
ctgaagctaa caccggcgca gacccgcttg gcgacggcgt tggagatccc aacacgatga 121920
cggactcaga ccctgaccct gatggcgacc cttgggatga cccgtgcctg gattcggtgg 121980
ccgtttcggt gtccgatccg tgcccttgtt ctgcaaccgc ccgcacttgc ggccctgatg 122040
ttgttggagc gggggcgcg ggtgcggcgg ctggcgcggc cgtgccacgc cccctgccgc 122100
gcctcccggc ctgttctcta gtgccgccgc gtcttccgca atgtcccctg caggaacttc 122160
gcaacccgtg tcccgatact atgccggaac gatgttgccg gctgccccac atgttcgcgg 122220
tttactgcgt gtaaaaaacg gcagattaag tagattgccc tttccacgac ccgcagccac 122280
ctgagcgcgc gcacgcagaa cctatgttta gtgcgattgc gattatgtcc gctaggtggc 122340
gggctatttt gatcctgccg cgcacgcatt gaccgttagg tggcgcagag cgccgttaac 122400
cgcagcgagc accgcacggc ggttttagtt tcgttactat aacgtttagt taaaatccgt 122460
ggccgatatt gggtggcagt gtgtaagtca cgcgcgccgc ggccccctta ccttgtcgcc 122520
ggcgaaccgc ggaaagtctg ttccgccgcg cagcgcgcgg ccgctgccga aaccaacgcc 122580
gcttgaattt cgcgggcggc tgggcttggt gtcagtgatt cacacagatg cgatgtgatt 122640
gggcgtggtg gtcaacgtca cttgttaacc cgtaaatctg gaaattcttg tccgccgcgc 122700
cggtgagccg cggccctaag ccttatagtg ctgtgcagcg atcctgaaac tcgaaagctg 122760
gccatataaa ccccagtatg gcgttgcgtg tcggtggcaa ccttttgaa aaagatttat 122820

tgccgccggg tgttaagcac caacaccgac cgtgcgtttt caattatgtc gggaggaatt 122880
acattaacgc tgctgctggc gacgcttgcc acggttcggt gcgctcttca aacgcactat 122940
gcggcggtcc ccgtgcactc taccgcgtcc ctggggtgcg tgttaacgac accccacgac 123000
gttcttatcg ttacctggca aaaacaggaa tcgcctagtc ccgttaacgt ggccacatat 123060
agttccgaag cgggcacggt ggttcagccc ccgttcgccg gtagggttga cattcccgaa 123120
cacaagttga ccagaacgac cctgaagttt tttaatgcca ccctggagga cgaggggtgc 123180
tacctgtgta tctttaacgc gtttggagtg ggaaagctgt cgggaaccgc ctgcttgacg 123240
gtttacgtcc ccctgtccat gtccgtcacg ttttaccccc cgattaaccc gacgcagctc 123300
gtctgtcggg ccgaggccag tcccgcaccg tcggtcaact ggaccggcgt gccgcccgag 123360
ctgtgcagcg aacctgaagt gtttccccgg cccaacggaa caaccctggt tgtgggtcgc 123420
tgcaacgtaa cgtcggtgga ccccgaagac cttgagaacg ccacgtgcct ggtcactcat 123480
ataggcggtt tggccgcggc gcggccccctg gaccccgtgt tttcggatcc cctggaaggg 123540
acgagccact acgtggtggg tgtggtggca gcggccgccg ttttaggcat tttttaacg 123600
ggtgtttttt tgtataggtc tatgtgagcg cgcgtgtccc ccgtgtctag tgttttgttc 123660
cccagtgagt gtctccatga caaatacaaa tttgaggctg gctttttagg gtggtttctt 123720
gtgcgacgct tcctgtgtaa ctgcatacac cggggtgtcg ccaggaaacc gcgtctcccc 123780
tttatgtccg ctcgccctcc cagagcgaaa gtgagaatgg ttcctggggc gttttggcgt 123840
tgagagagtc gggcgatgtt gccgtagcgg cgtctgcaaa ggctcacccg cttctgtttt 123900
ttttcttttt gtcagacaac aacatggacg ccttgaacaa taaccttaac ctgctgatgg 123960
atttctgtc taactattcg aatagctaca gtagttatga cgacaatatg tcttacacct 124020
tagacacgga atccacgctg tgtcggctga cggtggtttt cccacctacc gtttatgcga 124080
ttatatgctt ttttatttt tgcattacgc tgtttgggaa cgcgttggtg ctatatattt 124140
tttttaaatt taaagcgctc gccaactctg tggatgtact gatggctggg ttgtgttgta 124200
actccctgtt tctgtgcgcg tcgtttttgt tcagctggct gctgtacgtc gcgccacaga 124260
tgctcacgtc cgcgacgtgc aaggtggaaa tctttttctt ttacctgtac acgtactttg 124320
gcgtgtacat tgtggtgtgt atcagcctta tcaggtgcct gttagttgtg ttttcccgcc 124380
gcccgtgggt caagcacggg gcctccggct ttctctgcgt gtgtgtgtct ttaatcgtgg 124440
cgctggcgct gtctgccaac gcgagcctct ataggacggc cctgcgtcac ccagagacca 124500
gcgagtggat atgctacgaa gatgccgggg aagataccgt caactggaag ctgagaatca 124560
gaaccaccag cgcgatctgc gggtttttgg ttccgtttgg gctgatggtg ctcttttacg 124620
gacttacgtg gtgtatggtt aaaagcacga agctggccag aaagggagcc gttaggggtg 124680
taattgtgac ggtggtggtg ctgtttttaa tttttgcct gccctatcac ctgtgcaact 124740
tttttgacac cctgttgagg accggttttc tggccgaaac gtgctacctc agggacgtga 124800
tcagcgtggc catgcacata tgctccctgc tacagagcat gtatagcgcg ttcgtgccag 124860
tcgtgtattc tggtcttggg tctctgttta ggagaagggt tagggatacc tggtccgtgt 124920
ttaggtgttt ttccacttca ggtagtttat gagacactca cgcgacactt ggttggattg 124980
tttgtgtaca tttattttca ttttgtgtac atttattttc attaaagcga tctgacctgc 125040
agaccttacc tgacgtttac tgtctgtttc ttatgcacca gaggaacagg gactggaagg 125100
ccagcgccca cggggaaact gtcatgtccg caagctcggg cggcacgtac gcccactgcc 125160
aggggtggaa cgtcacggcg gggtcgacca gggaggccag gtggcgcccg tccggtgacg 125220
tgatggcggc cacggccgag ttggcagacg ggttgcgtgg gtagtgctgc gcgaacatcc 125280
tcgggtcggc gttgcccgtg tggtagttca gggcgatgcg ctgctgctgg ttgagatggt 125340
actccatggc gtcgcgcggg tatctcacgc ccaggtaccg gccgttggcc caccctggga 125400
ggacgaggcc ccgaagaacc ctaaacatga tgctgatggt ggtctggggg atgtggaggt 125460
ttagccagag gcactcgtgg ttccctgatg cgttctcctc caggtggatg tcccactggt 125520
cgggggtttt gggtccgggc gtgtcgtgag ggtctctct aagaagaccg agcgcccccca 125580
ggagctggaa cccaaactcc ccgcagcata gtgaaaatgt atccgctcgg cggaagaagg 125640
ccataaggcc cccatagcac ccagggtcgt tgagaagacc catgatcgcg catcgggccc 125700
ccacgtagct gtcttcgatg cccacggttc caccgatggt cagaccggag aatccccgga 125760
ggatgttccc tcctctaagg tcgtcggtgg agacggccgc gacgtcgaac ccgacgttgg 125820
tgaaggcggc catcagcgcc ctgggaagcg gggcaccggg ggtgaccaag gcggccactg 125880
ctggcggcct cgacggcgtt gcaaacagag tcagttcgct gtttctgcaa acctcggcga 125940
ggtggcccag gttgtgttgg ttgcacccgt agtcctttct gtagagttcc tgcgcgggcg 126000
tgaagctggg cccccatgag taccactgtt cgtccgagaa cgaggtccag tttgccgcca 126060
ccgaagtgag ggtctgtgaa tagacttcat cgttgttgtg tgagatgacg attctttctg 126120
acagaccctc ctgacccacg gtgccgcaca cggtggcccg gcagtcaaag ttttggcacg 126180
cctggcgcac ctcgtcgacg tgctggggct ggatttcgaa gatgacgccc ggggtttcgg 126240
acaccagcca ctgcaggggc gtttcctctg acgggatgcg aatctgtagg cctcggttac 126300
cggccagggc catctcgata gcggtaacca ccatgcctcc gtcgctgacg tcgtggccgg 126360
acaccaccag accccgggag agcagggcct ctgttagcat gaagaggttg gccaggtgtg 126420
tcgcctgaac gtccggtagg gtgggactcg gcaagagaca caggtgctca aaggtcgatc 126480

cctgggtcag gtgcggcacg gggaaggaca ccagcacgat gaggttaccg gtggccttca 126540
ggtcaggtgt gacgcgttgc ctggacgact tcacctcggc ggtggccgtg atgacgactg 126600
cgttaaacgg cacgggggct accgtctgat gctgttgggt ggcgctgatc agttcttccg 126660
acaggcacgg ggtgctgccg gccgacgtga tggcgaagtt gatgctcagg tccctgcata 126720
gttccttgca ggcaaacagt gtgtgctgca ggagccaggc ctggtcgtcc tccgggttcc 126780
aaccgaccga cgccgtcagc gtgatgtcgc ttagacgccg cacgtgcgct agcatgatgt 126840
tggtgaccgc ctcgcagatg gcgtatctcg cgcccactgc cgcgttgatg cccatcttgt 126900
atgcctgttc tccataagcc acgcactgtc cgacccaccg gttcctgccc ctgccctggg 126960
tgacccaggt ttctgggcta accaggcgtt ccgtcgcgtc ctgcctggtg atcggctggg 127020
ggcgatatat gggcacccgc gccattctgt ccgtaaacac cgacgtgtgg ttgacaatgt 127080
ggtaatctga cagcggccgg cccaggggtc ccacttcgca ctgctgagcc acgagtccgt 127140
tggagcatct gtccacgtgt cgcgtgacaa actctttgct gccaaccgtc ggacatcgca 127200
gaagctgatc gaccacggag gccaggcgga ataggcccca gtcggtcggc gagggtgcct 127260
gttcgccacc atcctcaggt tcttctgggc gcaccagcca gttggaggaa acggctgct 127320
cgtcgaaccc aaaagttgcg atgacttcgg ccgatgttgg ctgtctgatc ggccacgctc 127380
ccgggttcca cagctccagg tcgttcacga actgaattcc aagctcttcg caggtttccc 127440
caacgacggc gaaagggcat ccgcagatgc gtgcggccct cctgatggca tccaacgggc 127500
ccggttccgg ctccgtgtcg cgcgggagga cggcaaagac cacgcaggca gacacgttga 127560
gaaacttgtt aactaggatg tcgttcacgg tttcccgtga gatgtctggc gtggcggtca 127620
gttggctcag aatggttggt ggaagctgag agacaaagag gcgcaggccg ccccgcgaga 127680
ccagggctcc caggtggccg atgacagagg cgggttcatg ggcgcgatta atgctggtaa 127740
ataggggccc ttccaaaatc tcgtaagcca gcgacaacgt ttgggtgatg cgggtaatac 127800
ctgcggcgct ccccaggtaa agcgagcgct ggttgccatc gccgtcaacg ggttcaaaag 127860
acccaagaca gatgataggc tgactgaata aagacctcct gaggaatgtg tatgcttgag 127920
gcacgcacga taacagcgat gtggtggtaa taacgttcgc ctcgctggtg gtggcgctgt 127980
ggacggtttt aagaaagcca ccgaggacag gcgtgccggt gcgcgccagc tgctggcagt 128040
ggagtgtcgc tgttgtgtta atagctctgt ggcgtgcgcg agaaaatgcg cccactgacg 128100
gcgtagtggc gtaaaatcca cacagggccg atctgatttc gcccccgggc agcagcgtgg 128160
catggaacag aatgttttgc agcgcacctg cacctgtctg aatggcgctg tgcaggtgtg 128220
taaatgagtc agctatgtgg gttacgtttt cgcacgggag gcgctggaaa acgcagctgt 128280
agctcctgcg ggccatgacg gcgtatctgg acaccgtata gctgttgccg ttggaggcct 128340
gcacggccag cggtaatatg ttccgttgta acgggagcat caccgcggcg catattgggt 128400
ctctctcggg caacggtctc caggcggtgg attgcaatgt caccgggagc tggcgctgca 128460
cgcgcacatc atattctact agtgtgccaa gtgtcagggc ggccaaggtg cccgggggta 128520
cccgatcggc gacaaaccgt ctcgtctggc agctcagatg tctgcccacc tctaaccgca 128580
acgttctatc gaggccgctg tctctaaggc actgcactat ttcgctagat cttgtggtag 128640
gttggcgcat caggtctggg ccatagggca gggatatagt atgatggccc tggaggtgcg 128700
ccgaccgagt ttccagaggt gcataattaa ccagaggcgc cagaagatgc cagataacgc 128760
ccagcgctgg gtcgtcatcg tcgttgtcaa agttatcaga ggggcgtgtt acaagccaaa 128820
gtaaaaattc actaaagcac tgttctggca caagcgccaa aacactttct gcgtccgaat 128880
tgtcatgaat aaaagcctcc tcttctggac taagggctgc agcggcccag cgagggttcg 128940
tgcgttgggc catatttgct aagtctagcc ggcgcaaaag caagcgggtg ggactttaat 129000
agggcgtata ggggttgtta ggtatgcggt ggggagccgt ccatgttaca tggacttcaa 129060
tattaattta tgttagtttt gggcaccttg ggatgttaca cttataactt cccccatgtc 129120
acagtaatca cttgacagat ccccttctaa ctcataacaa ctgtttcggg gtgttgaggt 129180
tacaggtaag ggtgggaggg gaatgttagg tggcgaacca aagcataacg tcgcggggcg 129240
ggagtgggag cctgtaaccc cagagatgtc acttacgctt gttaggcgat cgacaactac 129300
ctccgaaagt gtgtcctgtt ggatgttttg tggattaggt aaaacagaaa gttgtggaga 129360
aagtaacctt atagatgacc tcctgctggc gcgtggctgg gtgacagctg gttgttgatg 129420
aatgaacact tcctcgggcg ttcctggggt ggtgggtgtg gtctcactaa ggacggccaa 129480
ctcaacacta acttcctccc gttcagaccg tgagcggttt cgtctgtagc gtcttgtgga 129540
gggggttgcc tgaacccttg ttggtgtaaa caatccaggt aaccggtaaa tgcacgcggt 129600
cagggcgatt tgtaacaaac ctaaaaacgc ataataagct tacaaacgca ttttggtaa 129660
cagtttgtat aatttagccc caaagtttat acaaagcgaa actacttacc gattagtaaa 129720
atgcaaaccg caagaaggat aacactcact gattctgttc caattggttc gaatttaaga 129780
atggagagta aatttgcccc gctgagaatg cctggaaata agaggcgcgt attttacatg 129840
tggtacaggg tagaaaacct tgtaggtttc agttgagtca ccctgccacg taaacccgcc 129900
tagagtgctg cacgcggtgt tggttgcagg aagtacatcc taattttaac tctaaaaggg 129960
gttccttgta aaattattta agttatggtg tacttacata atatccaatg agtagttaca 130020
ctccaaaaag atactcgata aaatgcgtga tatgtgaagt ttgcaacaaa tgcagcggtg 130080
agtaggccca agaaaatgat aaaccctata tgaggaaatt ttattagtac gtttgtgtaa 130140

[redacted]

ctgcaaaata aaattttaaa taaatacttt atattcatgt taccaagttt aatttgtggt 130200
ggttgttctt gaattatatg tactcctagt tctgaggcaa ctgtatgtag tgcaagcacg 130260
ctgccttgca acataaccaa taggcatgtt ccatgattcg atattcgcca tgccatttta 130320
agtgccacgc agattgcagc aatagctata aaaaaacaaa agtatgttaa tagtctataa 130380
agtataaaac attaaaacaa aattaaaaaa ttaaatacct gtatttgccg taactataga 130440
gtatgataat agtaaggtgt aagactgtgt tgttgtagtt tctgttaatg cataaagtcc 130500
cacaagtgct aaaataaaaa aataaatgtt agttatttaa tttttaaaag tttatactat 130560
tcaagtaatt caaagttgtg tcatacctgc gtaaattgca aaccaagatc caattattct 130620
tagtatgttt gtttttccac atagccacac ataaccataa aacacaatgg gcattaaaat 130680
tactggcaac agcgcagata ctttgtagtt tggagtccac ataagaaatg ccgtaagaat 130740
aacagtgcct atccatgttg gtcccacaag aagacatgat gttgttttac tagaaacaat 130800
tttcggtggt tttataacaa atattaatat gcctgaaaaa tataaacaaa cgattaattt 130860
aaggccaata tcatagtaaa acaaacttta aacaggtaaa aagttacgtc ttacctatgt 130920
aaaattgaga tacgcatgca atagcactac acgtctgtac caaaatactt gcgttgtcca 130980
aggccaatgc cgatagtaaa acacaacatg ctgcaaacaa gtacacaaaa tataagttta 131040
atttgctaat taaaattaat atttcaaatg tattataaaa tcatacctgt aatccatggg 131100
aaaaaaattg aaggcaacgg agttaattga cctgctctgc ttcgattaca atacaaacta 131160
atagttttcc ataaaagaga acaacacgta acaaatgaac aaataattgt gactgcttga 131220
ttagaagtgt ccctggcatg ttgtaccaag gctattatcc aaaacgttaa tgtaagaagg 131280
tttgtaagta accaactact ccaaataaaa aaggtatagc tacacgccat tttaaaataa 131340
tcaaaactgt attatcaagt taaggggggag tttgagcttt aaaaggttgg ttacatcgtt 131400
tccaggcaaa cctcacactg agtcacacaa aggtttctgt gtgcccctgt acgggaatct 131460
gccttattgc tgattttgtt ggcaaccaac caaagaaatc atggaaaagc agagtgggaa 131520
tatttgtttt tgttcgttgg cctggatgct tggaaacatt cctgttattt tgatgttagg 131580
ccaaactgtg gaaaacatta agagttgtta ttgaaaatat tttttggtaa caaaaccatg 131640
ttttaggtcc caaaaggaac ataaaaatgt ttaatagaaa catgcctagt acagaacatt 131700
ttgtttggta cgtaaactaa aaatatgtat ggtaagggct aggggtaagg gcaaggggta 131760
agggctaggg gtaagggcaa ggggtaaggg caagggtaa gggcaagggg taagggcaag 131820
gggtaagggc aaggggtaag ggcaaggggt aagggcaagg ggtaagggca aggggtaagg 131880
gctaggggta agggctaggg gtaagggcta ggggtaaggg ctagggtgag ctagggtgag 131940
ctagggttat agttatagtg taaaataata ggcctacgat aaaatgtcaa catatatttt 132000
ttaatttgtg tttcattaac agccatggtg tattagagta aaggttaacg cttcaataca 132060
atatacaagt aaaagtaaac cacacgaatt tataacatat ttacaaaagc gaaccaagct 132120
ggcacatgta acttcacgct accctaaacg tacaatactg ggactagaac ccagaggtag 132180
ttagaatata cggtagttac agaactttgc agttccctta ggccagcagg gctctgcggt 132240
taattaaaca aagtttaaag ataactgaaa ctttaggaag tgcgtatggg tgcaatgtgt 132300
tccaaatagg gcaagggtta cataaactgt tgcctagcgg ccgggcccgg aggcgcccgg 132360
ccggcgccgc cgccgggccg cggccgccat cttgcgcccg gggcgagggt cccccgcgcg 132420
cccccgggc ccgcgcgccg ggcgccgccg gcctcccccg gctcccggcc ctccgccccg 132480
ccggcccggg cccgcggccg ccgccccccgg ccgccccgcg ccgcgcgccg ggccgctttc 132540
ggttcgcggg gccgggggtc ccgcgggggg ccggggccgc ccccgggtgc cgctccgccg 132600
ggcccggccg actcccggga gcgcccccggt ccggccgagc ccggcggcgc cgcgccccga 132660
cgccccccgg gcccggggcc ccacaagccg cggcgcgcgg ctcccgatgc cgggcggccg 132720
ccgcccggca tggcggtcct ccgccggcct cccctccccc acgccgcccc gaaaggtggt 132780
ctccgcgccg ccgggagggg ggccggggcc cggggcgcct cggcggggcc cggcgcgggg 132840
cgcgaccgag ggcccccggga gaacggggga tcgggaaaac gcgaggggag cgggggacag 132900
gggacggcgt gtgcgtgctt gtgagacacc gggtacggct gcctgcctgc tcgctggcct 132960
gcttgctgag gggacagtag gcctgcttgc tcgctggcct gcttgctgag gggacagtag 133020
gcctgcttgc tgaggggaca gtaggcctgc ttgctcgctg gcctgcttgc tgaggggaca 133080
gtagggctgc tggcttgcta gtagggctgc tcgctggcct gcttgcttgc tcgctggcct 133140
gcttgcttgc tcgctggcct gcttgcttgc tcgctggcct gcttgcttgc tcgctggcct 133200
gcttgcttgc tcgctggcct gcttgcttgc tcgctggcct gcttgcttgc tcgctggcct 133260
gcttgctgag gggacagtag ggctgcttgc ttgctaaggg gacggtacgc ctgcctgatg 133320
gcttgatagt agggctgctg ggctgctagt agggctgctg ggctgctagt agggctgctg 133380
ggctgctagt agggctgctg ggctgctagt agggctcctg ggctgctagt agggctgctg 133440
ggctgctagt agggctcctg ggctgctagt agggctgctg ggctgctagt agggctgctg 133500
ggctgctagt agggctgctg ggctgctagt agggctgcct gctggcttgc ttgcttgctt 133560
gctagtgggg ccgcttgcct gctactaggg ctgctgtgca gctgggagaa cagagtaggg 133620
ctgccggcca gctgcgtgcg agggcgtccg agggccagac gaggacacgg gacccgggcc 133680
tctcccccgc ccggaccgcc gggcacccgg cccagatct 133719

```
<210> 2
<211> 1272
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1272)

<400> 2
atg ttt gtg ttg gtt tta ttt atg tta ttg caa cct gta tcg gtg gag      48
Met Phe Val Leu Val Leu Phe Met Leu Leu Gln Pro Val Ser Val Glu
1               5                   10                  15

ctt ttg cct gca aaa tta act tct gtt cca acg tgg tgt cca cca cat      96
Leu Leu Pro Ala Lys Leu Thr Ser Val Pro Thr Trp Cys Pro Pro His
            20                  25                  30

ccc gga gat acc tac ttg cta acc tgc cgc ggg acg tct acg gcc aga      144
Pro Gly Asp Thr Tyr Leu Leu Thr Cys Arg Gly Thr Ser Thr Ala Arg
        35                  40                  45

gac cag cga agc aca caa tgg ttt cgc aac aac acg ctt atg cgt ggg      192
Asp Gln Arg Ser Thr Gln Trp Phe Arg Asn Asn Thr Leu Met Arg Gly
    50                  55                  60

agt aat ttc tac ggc aga ctg gta tct gtg act ccc aat gct acg ata      240
Ser Asn Phe Tyr Gly Arg Leu Val Ser Val Thr Pro Asn Ala Thr Ile
65                  70                  75                  80

tct gac cgg tat gcg tgt caa aca aaa aca aca acg cgg agt aac aac      288
Ser Asp Arg Tyr Ala Cys Gln Thr Lys Thr Thr Thr Arg Ser Asn Asn
                85                  90                  95

atc gat ttt cgg gta agc tca tcg cgc ctc acg ctc caa gaa cgg tgc      336
Ile Asp Phe Arg Val Ser Ser Ser Arg Leu Thr Leu Gln Glu Arg Cys
            100                 105                 110

tct tca tac ggc tat act tac gcg aat aac aca agg gta ttg agg tgt      384
Ser Ser Tyr Gly Tyr Thr Tyr Ala Asn Asn Thr Arg Val Leu Arg Cys
        115                 120                 125

tac tct ggt gga aac gta act tta aga aac gtt gtc ttt cat tta aac      432
Tyr Ser Gly Gly Asn Val Thr Leu Arg Asn Val Val Phe His Leu Asn
    130                 135                 140

ggt aca gcg gtc atc aac ggt act aca aca aac ata cat aca ttt gtg      480
Gly Thr Ala Val Ile Asn Gly Thr Thr Thr Asn Ile His Thr Phe Val
145                 150                 155                 160

tta aca gaa aag aca gga ggg acg tat ttc tgt tct gcg ttt att ggg      528
Leu Thr Glu Lys Thr Gly Gly Thr Tyr Phe Cys Ser Ala Phe Ile Gly
                165                 170                 175

aat gaa aaa ttc tat tct cag aca att aat gtg ttt ttt act tca ttt      576
Asn Glu Lys Phe Tyr Ser Gln Thr Ile Asn Val Phe Phe Thr Ser Phe
            180                 185                 190

acc ttt aaa cct aca aac gac att ccc aat gag tca cat ttt aat aaa      624
Thr Phe Lys Pro Thr Asn Asp Ile Pro Asn Glu Ser His Phe Asn Lys
        195                 200                 205
```

```
act ggg caa ata caa caa aca gct agt gta caa cat cct gaa aac tac    672
Thr Gly Gln Ile Gln Gln Thr Ala Ser Val Gln His Pro Glu Asn Tyr
    210                 215                 220

gtt gtg ttc tct gtt ccc gtt ttt tct att ggc gtt tta aca ggt att    720
Val Val Phe Ser Val Pro Val Phe Ser Ile Gly Val Leu Thr Gly Ile
225                 230                 235                 240

gca ata tcg ttg att atg tgt tgg tta ttt aca ata cgc tgc aac gag    768
Ala Ile Ser Leu Ile Met Cys Trp Leu Phe Thr Ile Arg Cys Asn Glu
                245                 250                 255

aac tct gaa tca tca act aat agt tat gca agc cag aca agc tac att    816
Asn Ser Glu Ser Ser Thr Asn Ser Tyr Ala Ser Gln Thr Ser Tyr Ile
            260                 265                 270

caa ccc tcc cat aat cag cgt tcc aat act aat gaa tgt agt cgc cat    864
Gln Pro Ser His Asn Gln Arg Ser Asn Thr Asn Glu Cys Ser Arg His
        275                 280                 285

acc tac aga aat gct cat caa gaa gag agt att gaa gaa cta cca aac    912
Thr Tyr Arg Asn Ala His Gln Glu Glu Ser Ile Glu Glu Leu Pro Asn
    290                 295                 300

caa cac aca agt gaa act gat tct tgc tgt caa tta gtt tta ctt gaa    960
Gln His Thr Ser Glu Thr Asp Ser Cys Cys Gln Leu Val Leu Leu Glu
305                 310                 315                 320

gtg aaa aat gta gcc tac gat gga ccg cag gaa aac aca att aac gaa    1008
Val Lys Asn Val Ala Tyr Asp Gly Pro Gln Glu Asn Thr Ile Asn Glu
                325                 330                 335

gtt atg gaa cag tat gat gat gtg gtt gta aaa aat ata gaa caa aca    1056
Val Met Glu Gln Tyr Asp Asp Val Val Val Lys Asn Ile Glu Gln Thr
            340                 345                 350

tca tat gag gat aat gtt gag cac atg gac tat agt gat act ata aat    1104
Ser Tyr Glu Asp Asn Val Glu His Met Asp Tyr Ser Asp Thr Ile Asn
        355                 360                 365

ccc aat ttt aat tac tac agt gga cta ata ttg gaa gaa gta gat gaa    1152
Pro Asn Phe Asn Tyr Tyr Ser Gly Leu Ile Leu Glu Glu Val Asp Glu
    370                 375                 380

gtt ttt tac aat gaa cta gaa aat caa tat cat gga tta ata ctg gag    1200
Val Phe Tyr Asn Glu Leu Glu Asn Gln Tyr His Gly Leu Ile Leu Glu
385                 390                 395                 400

aat tta gat cac aat gag tac aat cat tta aat gaa tta aac atg ata    1248
Asn Leu Asp His Asn Glu Tyr Asn His Leu Asn Glu Leu Asn Met Ile
                405                 410                 415

gaa caa tat gat tgg tta gaa taa                                    1272
Glu Gln Tyr Asp Trp Leu Glu
            420
```

<210> 3
<211> 423
<212> PRT

<213> Macaca mulatta rhadinovirus 17577

<400> 3

Met Phe Val Leu Val Leu Phe Met Leu Leu Gln Pro Val Ser Val Glu
1 5 10 15
Leu Leu Pro Ala Lys Leu Thr Ser Val Pro Thr Trp Cys Pro Pro His
20 25 30
Pro Gly Asp Thr Tyr Leu Leu Thr Cys Arg Gly Thr Ser Thr Ala Arg
35 40 45
Asp Gln Arg Ser Thr Gln Trp Phe Arg Asn Asn Thr Leu Met Arg Gly
50 55 60
Ser Asn Phe Tyr Gly Arg Leu Val Ser Val Thr Pro Asn Ala Thr Ile
65 70 75 80
Ser Asp Arg Tyr Ala Cys Gln Thr Lys Thr Thr Thr Arg Ser Asn Asn
85 90 95
Ile Asp Phe Arg Val Ser Ser Ser Arg Leu Thr Leu Gln Glu Arg Cys
100 105 110
Ser Ser Tyr Gly Tyr Thr Tyr Ala Asn Asn Thr Arg Val Leu Arg Cys
115 120 125
Tyr Ser Gly Gly Asn Val Thr Leu Arg Asn Val Val Phe His Leu Asn
130 135 140
Gly Thr Ala Val Ile Asn Gly Thr Thr Thr Asn Ile His Thr Phe Val
145 150 155 160
Leu Thr Glu Lys Thr Gly Gly Thr Tyr Phe Cys Ser Ala Phe Ile Gly
165 170 175
Asn Glu Lys Phe Tyr Ser Gln Thr Ile Asn Val Phe Phe Thr Ser Phe
180 185 190
Thr Phe Lys Pro Thr Asn Asp Ile Pro Asn Glu Ser His Phe Asn Lys
195 200 205
Thr Gly Gln Ile Gln Gln Thr Ala Ser Val Gln His Pro Glu Asn Tyr
210 215 220
Val Val Phe Ser Val Pro Val Phe Ser Ile Gly Val Leu Thr Gly Ile
225 230 235 240
Ala Ile Ser Leu Ile Met Cys Trp Leu Phe Thr Ile Arg Cys Asn Glu
245 250 255
Asn Ser Glu Ser Ser Thr Asn Ser Tyr Ala Ser Gln Thr Ser Tyr Ile
260 265 270
Gln Pro Ser His Asn Gln Arg Ser Asn Thr Asn Glu Cys Ser Arg His
275 280 285
Thr Tyr Arg Asn Ala His Gln Glu Glu Ser Ile Glu Glu Leu Pro Asn
290 295 300
Gln His Thr Ser Glu Thr Asp Ser Cys Cys Gln Leu Val Leu Leu Glu
305 310 315 320
Val Lys Asn Val Ala Tyr Asp Gly Pro Gln Glu Asn Thr Ile Asn Glu
325 330 335
Val Met Glu Gln Tyr Asp Asp Val Val Val Lys Asn Ile Glu Gln Thr
340 345 350
Ser Tyr Glu Asp Asn Val Glu His Met Asp Tyr Ser Asp Thr Ile Asn
355 360 365
Pro Asn Phe Asn Tyr Tyr Ser Gly Leu Ile Leu Glu Glu Val Asp Glu
370 375 380
Val Phe Tyr Asn Glu Leu Glu Asn Gln Tyr His Gly Leu Ile Leu Glu
385 390 395 400
Asn Leu Asp His Asn Glu Tyr Asn His Leu Asn Glu Leu Asn Met Ile
405 410 415
Glu Gln Tyr Asp Trp Leu Glu
420

<210> 4

```
<211> 567
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(567)

<400> 4
atg gac att aca gtt aac tgc atc gtt gca gtt gat gaa caa tta gga      48
Met Asp Ile Thr Val Asn Cys Ile Val Ala Val Asp Glu Gln Leu Gly
  1               5                  10                  15

att ggt aaa aac gga aca atg cct tgg ccg tac ctt aga aac gaa atg      96
Ile Gly Lys Asn Gly Thr Met Pro Trp Pro Tyr Leu Arg Asn Glu Met
             20                  25                  30

atg tat ttt cag aaa atg acg tct acg cca tcg gtt gta ggt gaa aag     144
Met Tyr Phe Gln Lys Met Thr Ser Thr Pro Ser Val Val Gly Glu Lys
         35                  40                  45

aat gtg gtt att atg ggg aaa aga aca tgg ttt tcc atc cct gaa aag     192
Asn Val Val Ile Met Gly Lys Arg Thr Trp Phe Ser Ile Pro Glu Lys
     50                  55                  60

aag cgc cct ctg gtg aac aga att aat att att ttg agc aga gaa tta     240
Lys Arg Pro Leu Val Asn Arg Ile Asn Ile Ile Leu Ser Arg Glu Leu
 65                  70                  75                  80

cgt gaa cca ccg cac gga gca cac ttt cta gcg aga acg cta gac gat     288
Arg Glu Pro Pro His Gly Ala His Phe Leu Ala Arg Thr Leu Asp Asp
                 85                  90                  95

gcg ttt aat ttc tat aga caa tac aaa cta aaa gaa caa tta aac acg     336
Ala Phe Asn Phe Tyr Arg Gln Tyr Lys Leu Lys Glu Gln Leu Asn Thr
            100                 105                 110

gtc tgg gtt att gga gga aaa tct gtt tat gag agt gtt ctt aac tat     384
Val Trp Val Ile Gly Gly Lys Ser Val Tyr Glu Ser Val Leu Asn Tyr
        115                 120                 125

aag tgt ccc tta aaa cta tat att acg cgc att atg gag agt ttt gat     432
Lys Cys Pro Leu Lys Leu Tyr Ile Thr Arg Ile Met Glu Ser Phe Asp
    130                 135                 140

tgt gat gta ttt ttt cca tca att aat ttt acc gaa tac acc atg tta     480
Cys Asp Val Phe Phe Pro Ser Ile Asn Phe Thr Glu Tyr Thr Met Leu
145                 150                 155                 160

tca gag ata ccg ggc aaa gat aca aac ttt gag gaa aat gga att aaa     528
Ser Glu Ile Pro Gly Lys Asp Thr Asn Phe Glu Glu Asn Gly Ile Lys
                165                 170                 175

tat aag ttt cag gta tat gaa aaa aat ttt aat aaa taa                 567
Tyr Lys Phe Gln Val Tyr Glu Lys Asn Phe Asn Lys
            180                 185

<210> 5
<211> 188
<212> PRT
```

<213> Macaca mulatta rhadinovirus 17577

<400> 5
Met Asp Ile Thr Val Asn Cys Ile Val Ala Val Asp Glu Gln Leu Gly
1   5   10   15
Ile Gly Lys Asn Gly Thr Met Pro Trp Pro Tyr Leu Arg Asn Glu Met
            20   25   30
Met Tyr Phe Gln Lys Met Thr Ser Thr Pro Ser Val Val Gly Glu Lys
        35   40   45
Asn Val Val Ile Met Gly Lys Arg Thr Trp Phe Ser Ile Pro Glu Lys
    50   55   60
Lys Arg Pro Leu Val Asn Arg Ile Asn Ile Ile Leu Ser Arg Glu Leu
65   70   75   80
Arg Glu Pro Pro His Gly Ala His Phe Leu Ala Arg Thr Leu Asp Asp
                85   90   95
Ala Phe Asn Phe Tyr Arg Gln Tyr Lys Leu Lys Glu Gln Leu Asn Thr
            100   105   110
Val Trp Val Ile Gly Gly Lys Ser Val Tyr Glu Ser Val Leu Asn Tyr
        115   120   125
Lys Cys Pro Leu Lys Leu Tyr Ile Thr Arg Ile Met Glu Ser Phe Asp
    130   135   140
Cys Asp Val Phe Phe Pro Ser Ile Asn Phe Thr Glu Tyr Thr Met Leu
145   150   155   160
Ser Glu Ile Pro Gly Lys Asp Thr Asn Phe Glu Glu Asn Gly Ile Lys
                165   170   175
Tyr Lys Phe Gln Val Tyr Glu Lys Asn Phe Asn Lys
            180   185

<210> 6
<211> 1938
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1938)

<400> 6
atg acg ttt aaa ctt ttt cct ctg ttt tta tta cac gcc ata atg tac 48
Met Thr Phe Lys Leu Phe Pro Leu Phe Leu Leu His Ala Ile Met Tyr
1   5   10   15

gtc cac tgc gat gaa aac tgt aaa cct cca cat ttc acg gaa tat cgc 96
Val His Cys Asp Glu Asn Cys Lys Pro Pro His Phe Thr Glu Tyr Arg
            20   25   30

gtc aag tct aac aca gaa aag gac tta tat agt gtt gga gaa aca gct 144
Val Lys Ser Asn Thr Glu Lys Asp Leu Tyr Ser Val Gly Glu Thr Ala
        35   40   45

gaa tta att tgt cgt cct ggt tat gtt aca aat aca aaa ata ata aca 192
Glu Leu Ile Cys Arg Pro Gly Tyr Val Thr Asn Thr Lys Ile Ile Thr
    50   55   60

aca gaa tgt tta caa aat ggt acg tgg tca aca cca aat ttt cca tgc 240
Thr Glu Cys Leu Gln Asn Gly Thr Trp Ser Thr Pro Asn Phe Pro Cys
65   70   75   80

gac aga aaa aga tgt ccc aca cct gct gac ttg ctg aat gga gcc gtg 288

Asp Arg Lys Arg Cys Pro Thr Pro Ala Asp Leu Leu Asn Gly Ala Val
85 90 95

cac att cac ggg gga gat aat gcc tta aaa ttt gga tcc aat att tcc 336
His Ile His Gly Gly Asp Asn Ala Leu Lys Phe Gly Ser Asn Ile Ser
100 105 110

tat gag tgt aat gaa ggt tat gat tta att ggt agt aat gtt cgt ttt 384
Tyr Glu Cys Asn Glu Gly Tyr Asp Leu Ile Gly Ser Asn Val Arg Phe
115 120 125

tgt att tta caa gac aca gaa aac gta aat tgg gat tca aat gaa cca 432
Cys Ile Leu Gln Asp Thr Glu Asn Val Asn Trp Asp Ser Asn Glu Pro
130 135 140

gtc tgt gaa att cag aaa tgt att aaa cca ccg gca gtg gaa cat ggg 480
Val Cys Glu Ile Gln Lys Cys Ile Lys Pro Pro Ala Val Glu His Gly
145 150 155 160

gac tac cta cct aac caa gat gtt tat aac tat gga gat gca att aca 528
Asp Tyr Leu Pro Asn Gln Asp Val Tyr Asn Tyr Gly Asp Ala Ile Thr
165 170 175

ttt aaa tgt tca ttg tcg tat aca ctc gtt gga tca aca aca tta gta 576
Phe Lys Cys Ser Leu Ser Tyr Thr Leu Val Gly Ser Thr Thr Leu Val
180 185 190

tgc acg tca aac aaa aag tgg tca aac tct ttc cca acg tgt tta atg 624
Cys Thr Ser Asn Lys Lys Trp Ser Asn Ser Phe Pro Thr Cys Leu Met
195 200 205

ctc gta tgt gaa agt cca caa ata gac aat ggg tac ata gac att ggc 672
Leu Val Cys Glu Ser Pro Gln Ile Asp Asn Gly Tyr Ile Asp Ile Gly
210 215 220

ttg tcc aga aga tac aac cat gga caa tca att act gta aag tgt agc 720
Leu Ser Arg Arg Tyr Asn His Gly Gln Ser Ile Thr Val Lys Cys Ser
225 230 235 240

gac ggg tac aac att gtt ggg cct gaa aca tta acg tgc aca aac aca 768
Asp Gly Tyr Asn Ile Val Gly Pro Glu Thr Leu Thr Cys Thr Asn Thr
245 250 255

act tgg gtt cca cca tta cct aag tgt gta tta gtt aca aat aac cca 816
Thr Trp Val Pro Pro Leu Pro Lys Cys Val Leu Val Thr Asn Asn Pro
260 265 270

agc aca ccc atg cca gaa aca ccc atg cca gaa aca ccc acg cca gat 864
Ser Thr Pro Met Pro Glu Thr Pro Met Pro Glu Thr Pro Thr Pro Asp
275 280 285

tat caa aaa ata aat ttg tca acc gct aaa act gca aca aca cca aat 912
Tyr Gln Lys Ile Asn Leu Ser Thr Ala Lys Thr Ala Thr Thr Pro Asn
290 295 300

gcg ttt gtt aca act gtt gtt tct cca gag aaa gac gac gtt act tgt 960
Ala Phe Val Thr Thr Val Val Ser Pro Glu Lys Asp Asp Val Thr Cys
305 310 315 320

gta aag cct cat ttt gag cga ttc atg gta aag gct gaa aat gac aag 1008
Val Lys Pro His Phe Glu Arg Phe Met Val Lys Ala Glu Asn Asp Lys

325 330 335

gaa aaa tac agt gtt ggt gcg agc gtt gag cta ata tgt cga cca gga 1056
Glu Lys Tyr Ser Val Gly Ala Ser Val Glu Leu Ile Cys Arg Pro Gly
340 345 350

ttt act aaa atg cag tct aca gtt tct gtt gaa tgt ttg tcc aac gga 1104
Phe Thr Lys Met Gln Ser Thr Val Ser Val Glu Cys Leu Ser Asn Gly
355 360 365

aca tgg act gct cca aat gcc aag tgt cat aga aaa aaa tgt cca acc 1152
Thr Trp Thr Ala Pro Asn Ala Lys Cys His Arg Lys Lys Cys Pro Thr
370 375 380

cct caa gaa ctt tta aac gga gag tat ata gtt aca agc gga gaa gat 1200
Pro Gln Glu Leu Leu Asn Gly Glu Tyr Ile Val Thr Ser Gly Glu Asp
385 390 395 400

gct ttt aag tac gga aca aat ata aca tat aaa tgt aat gaa ggt tat 1248
Ala Phe Lys Tyr Gly Thr Asn Ile Thr Tyr Lys Cys Asn Glu Gly Tyr
405 410 415

caa ctt tta gga agt atg gtg cgg att tgt atg ctt aaa gac gat tta 1296
Gln Leu Leu Gly Ser Met Val Arg Ile Cys Met Leu Lys Asp Asp Leu
420 425 430

aaa aca gtt gac tgg gag cca aaa gcg cct ata tgt gat att gaa aaa 1344
Lys Thr Val Asp Trp Glu Pro Lys Ala Pro Ile Cys Asp Ile Glu Lys
435 440 445

tgt aag cca ccg cca caa att aca aac gga aaa tac cat ccg gtg aaa 1392
Cys Lys Pro Pro Pro Gln Ile Thr Asn Gly Lys Tyr His Pro Val Lys
450 455 460

gac ttt tat cag tat ttg gac acc gta aca ttt tcg tgc aat cgt gac 1440
Asp Phe Tyr Gln Tyr Leu Asp Thr Val Thr Phe Ser Cys Asn Arg Asp
465 470 475 480

ttt tct tta gtt gga gat gaa atg aca acg tgt ata agt aat acg tgg 1488
Phe Ser Leu Val Gly Asp Glu Met Thr Thr Cys Ile Ser Asn Thr Trp
485 490 495

aat aaa ccg ttt cca aga tgt gaa caa atc act tgc agc gct cct aat 1536
Asn Lys Pro Phe Pro Arg Cys Glu Gln Ile Thr Cys Ser Ala Pro Asn
500 505 510

att gca cac gga aag ctg cta aca ggt tct tca agc gtt tac aaa tac 1584
Ile Ala His Gly Lys Leu Leu Thr Gly Ser Ser Ser Val Tyr Lys Tyr
515 520 525

ggt caa tct gtt acc att ggt tgt gaa act gga ttt act cta att ggc 1632
Gly Gln Ser Val Thr Ile Gly Cys Glu Thr Gly Phe Thr Leu Ile Gly
530 535 540

agt gaa att tct aca tgc aag gat tca tcg tgg gat cca cca ctt cct 1680
Ser Glu Ile Ser Thr Cys Lys Asp Ser Ser Trp Asp Pro Pro Leu Pro
545 550 555 560

acg tgc gtg cca gct gtt tca atg cct tct gac aca cct aaa cca gaa 1728
Thr Cys Val Pro Ala Val Ser Met Pro Ser Asp Thr Pro Lys Pro Glu
565 570 575

```
acc aaa aaa cca aac acg cca acg cca gaa gca ccc aaa cca aac acc      1776
Thr Lys Lys Pro Asn Thr Pro Thr Pro Glu Ala Pro Lys Pro Asn Thr
            580                 585                 590

cca aac gtt gga aca cat aca cca ttc aaa cca cca cca caa aat cca      1824
Pro Asn Val Gly Thr His Thr Pro Phe Lys Pro Pro Pro Gln Asn Pro
        595                 600                 605

cca ata gca ccc cca atg agt aaa tgg aaa agg cat gtc gtg tta gtt      1872
Pro Ile Ala Pro Pro Met Ser Lys Trp Lys Arg His Val Val Leu Val
    610                 615                 620

ctt ttt gca agt gtc gcg tcc ttg tta ttc gta ctt gct gcc ctt tat      1920
Leu Phe Ala Ser Val Ala Ser Leu Leu Phe Val Leu Ala Ala Leu Tyr
625                 630                 635                 640

tgt tgt ttt cta aaa taa                                              1938
Cys Cys Phe Leu Lys
                645
```

<210> 7
<211> 645
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 7

```
Met Thr Phe Lys Leu Phe Pro Leu Phe Leu Leu His Ala Ile Met Tyr
 1               5                  10                  15
Val His Cys Asp Glu Asn Cys Lys Pro Pro His Phe Thr Glu Tyr Arg
            20                  25                  30
Val Lys Ser Asn Thr Glu Lys Asp Leu Tyr Ser Val Gly Glu Thr Ala
        35                  40                  45
Glu Leu Ile Cys Arg Pro Gly Tyr Val Thr Asn Thr Lys Ile Ile Thr
    50                  55                  60
Thr Glu Cys Leu Gln Asn Gly Thr Trp Ser Thr Pro Asn Phe Pro Cys
65                  70                  75                  80
Asp Arg Lys Arg Cys Pro Thr Pro Ala Asp Leu Leu Asn Gly Ala Val
                85                  90                  95
His Ile His Gly Gly Asp Asn Ala Leu Lys Phe Gly Ser Asn Ile Ser
            100                 105                 110
Tyr Glu Cys Asn Glu Gly Tyr Asp Leu Ile Gly Ser Asn Val Arg Phe
        115                 120                 125
Cys Ile Leu Gln Asp Thr Glu Asn Val Asn Trp Asp Ser Asn Glu Pro
    130                 135                 140
Val Cys Glu Ile Gln Lys Cys Ile Lys Pro Pro Ala Val Glu His Gly
145                 150                 155                 160
Asp Tyr Leu Pro Asn Gln Asp Val Tyr Asn Tyr Gly Asp Ala Ile Thr
                165                 170                 175
Phe Lys Cys Ser Leu Ser Tyr Thr Leu Val Gly Ser Thr Thr Leu Val
            180                 185                 190
Cys Thr Ser Asn Lys Lys Trp Ser Asn Ser Phe Pro Thr Cys Leu Met
        195                 200                 205
Leu Val Cys Glu Ser Pro Gln Ile Asp Asn Gly Tyr Ile Asp Ile Gly
    210                 215                 220
Leu Ser Arg Arg Tyr Asn His Gly Gln Ser Ile Thr Val Lys Cys Ser
225                 230                 235                 240
Asp Gly Tyr Asn Ile Val Gly Pro Glu Thr Leu Thr Cys Thr Asn Thr
                245                 250                 255
Thr Trp Val Pro Pro Leu Pro Lys Cys Val Leu Val Thr Asn Asn Pro
```

260 265 270
Ser Thr Pro Met Pro Glu Thr Pro Met Pro Glu Thr Pro Thr Pro Asp
275 280 285
Tyr Gln Lys Ile Asn Leu Ser Thr Ala Lys Thr Ala Thr Thr Pro Asn
290 295 300
Ala Phe Val Thr Thr Val Val Ser Pro Glu Lys Asp Asp Val Thr Cys
305 310 315 320
Val Lys Pro His Phe Glu Arg Phe Met Val Lys Ala Glu Asn Asp Lys
325 330 335
Glu Lys Tyr Ser Val Gly Ala Ser Val Glu Leu Ile Cys Arg Pro Gly
340 345 350
Phe Thr Lys Met Gln Ser Thr Val Ser Val Glu Cys Leu Ser Asn Gly
355 360 365
Thr Trp Thr Ala Pro Asn Ala Lys Cys His Arg Lys Lys Cys Pro Thr
370 375 380
Pro Gln Glu Leu Leu Asn Gly Glu Tyr Ile Val Thr Ser Gly Glu Asp
385 390 395 400
Ala Phe Lys Tyr Gly Thr Asn Ile Thr Tyr Lys Cys Asn Glu Gly Tyr
405 410 415
Gln Leu Leu Gly Ser Met Val Arg Ile Cys Met Leu Lys Asp Asp Leu
420 425 430
Lys Thr Val Asp Trp Glu Pro Lys Ala Pro Ile Cys Asp Ile Glu Lys
435 440 445
Cys Lys Pro Pro Pro Gln Ile Thr Asn Gly Lys Tyr His Pro Val Lys
450 455 460
Asp Phe Tyr Gln Tyr Leu Asp Thr Val Thr Phe Ser Cys Asn Arg Asp
465 470 475 480
Phe Ser Leu Val Gly Asp Glu Met Thr Thr Cys Ile Ser Asn Thr Trp
485 490 495
Asn Lys Pro Phe Pro Arg Cys Glu Gln Ile Thr Cys Ser Ala Pro Asn
500 505 510
Ile Ala His Gly Lys Leu Leu Thr Gly Ser Ser Ser Val Tyr Lys Tyr
515 520 525
Gly Gln Ser Val Thr Ile Gly Cys Glu Thr Gly Phe Thr Leu Ile Gly
530 535 540
Ser Glu Ile Ser Thr Cys Lys Asp Ser Ser Trp Asp Pro Pro Leu Pro
545 550 555 560
Thr Cys Val Pro Ala Val Ser Met Pro Ser Asp Thr Pro Lys Pro Glu
565 570 575
Thr Lys Lys Pro Asn Thr Pro Thr Pro Glu Ala Pro Lys Pro Asn Thr
580 585 590
Pro Asn Val Gly Thr His Thr Pro Phe Lys Pro Pro Pro Gln Asn Pro
595 600 605
Pro Ile Ala Pro Pro Met Ser Lys Trp Lys Arg His Val Val Leu Val
610 615 620
Leu Phe Ala Ser Val Ala Ser Leu Leu Phe Val Leu Ala Ala Leu Tyr
625 630 635 640
Cys Cys Phe Leu Lys
645

<210> 8
<211> 3399
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(3399)

<400> 8

```
atg gct tcc aaa ggc aac gcc gga caa ccc ctg gaa gat aat cag ggg    48
Met Ala Ser Lys Gly Asn Ala Gly Gln Pro Leu Glu Asp Asn Gln Gly
  1               5                  10                  15

tct cgt gcc ccg ata ggt gcg tgc gga tac gtg tac gcg tat tcg aaa    96
Ser Arg Ala Pro Ile Gly Ala Cys Gly Tyr Val Tyr Ala Tyr Ser Lys
             20                  25                  30

caa gac ttt ccc ttt gcc gag gcg tcc ata ctc ggc aac aga cca tct    144
Gln Asp Phe Pro Phe Ala Glu Ala Ser Ile Leu Gly Asn Arg Pro Ser
         35                  40                  45

gga tct ggc gtt ttc tcg cta cca atc ctt tac gga ctt aca gtt gaa    192
Gly Ser Gly Val Phe Ser Leu Pro Ile Leu Tyr Gly Leu Thr Val Glu
     50                  55                  60

cac gaa ttc cct ctc acc gta aaa gcc gca tac aaa aaa gtt gac acc    240
His Glu Phe Pro Leu Thr Val Lys Ala Ala Tyr Lys Lys Val Asp Thr
 65                  70                  75                  80

acg acg ctc gcc gtt aag gtg acg tgc ttt cac aga gag gtt att gtg    288
Thr Thr Leu Ala Val Lys Val Thr Cys Phe His Arg Glu Val Ile Val
                 85                  90                  95

ttt cac aat gca agt tta ttc agg ccg gtg ttt gac ggt acc ggt ctt    336
Phe His Asn Ala Ser Leu Phe Arg Pro Val Phe Asp Gly Thr Gly Leu
            100                 105                 110

aac gaa cta tgc gag gaa gcc agg gct ctc ttt ggg tac acg cag ttt    384
Asn Glu Leu Cys Glu Glu Ala Arg Ala Leu Phe Gly Tyr Thr Gln Phe
        115                 120                 125

ata gaa ccg ggt cca cct cac agc ata tgg aac cct ctg gaa tgt ccg    432
Ile Glu Pro Gly Pro Pro His Ser Ile Trp Asn Pro Leu Glu Cys Pro
    130                 135                 140

cag tta ccg gac aag gat gag atg ttt ctt ggc gtt gtt gtt acg gaa    480
Gln Leu Pro Asp Lys Asp Glu Met Phe Leu Gly Val Val Val Thr Glu
145                 150                 155                 160

ggg ttt aag gaa aga ctg tgg agg ggc tgt ctc gtt ccc gcg gtg ttc    528
Gly Phe Lys Glu Arg Leu Trp Arg Gly Cys Leu Val Pro Ala Val Phe
                165                 170                 175

cag acc cag cag gtg cag att gcc gga cgc cag gcg ttt aaa gtg ccg    576
Gln Thr Gln Gln Val Gln Ile Ala Gly Arg Gln Ala Phe Lys Val Pro
            180                 185                 190

ttg tac gac gaa gac ctg ttt gca cct cac ggt cat aga atg cca agg    624
Leu Tyr Asp Glu Asp Leu Phe Ala Pro His Gly His Arg Met Pro Arg
        195                 200                 205

ttt tac cat aaa gac gtt agc gcg tac ctc tac gac tcc ctc ttt acc    672
Phe Tyr His Lys Asp Val Ser Ala Tyr Leu Tyr Asp Ser Leu Phe Thr
    210                 215                 220

agc atc gcc cag gcc ctg aga ctc aaa gac gtg acg gcg gtc atc cac    720
Ser Ile Ala Gln Ala Leu Arg Leu Lys Asp Val Thr Ala Val Ile His
225                 230                 235                 240
```

gcc aca gaa aag caa ttc atg cag gac cat tac aaa att gcc aag ata 768
Ala Thr Glu Lys Gln Phe Met Gln Asp His Tyr Lys Ile Ala Lys Ile
245 250 255

gtg cag gca aaa cag ttt tca acg acg ctg ccg aaa acg aca gac ggg 816
Val Gln Ala Lys Gln Phe Ser Thr Thr Leu Pro Lys Thr Thr Asp Gly
260 265 270

tcg tcc cac atg att gtg gac agc gtc gtc gcc gag ctc gcc ctt agt 864
Ser Ser His Met Ile Val Asp Ser Val Val Ala Glu Leu Ala Leu Ser
275 280 285

tac ggc tgt atg ttt ctc gag tgt ccc cag gac gcg tgc gag ttg ctg 912
Tyr Gly Cys Met Phe Leu Glu Cys Pro Gln Asp Ala Cys Glu Leu Leu
290 295 300

aac tac gat agc tgg ccc ata ttt gat ggt tgt gac tca cca gag gct 960
Asn Tyr Asp Ser Trp Pro Ile Phe Asp Gly Cys Asp Ser Pro Glu Ala
305 310 315 320

agg gtt aac gcg tta gag cgc tgg tcg gcc gaa cag gcc gtt cac gtg 1008
Arg Val Asn Ala Leu Glu Arg Trp Ser Ala Glu Gln Ala Val His Val
325 330 335

gcg ggt caa ctg ttc gct gcc aat tcg gtg ctg tac cta act aaa gtg 1056
Ala Gly Gln Leu Phe Ala Ala Asn Ser Val Leu Tyr Leu Thr Lys Val
340 345 350

cag aag caa gcg ccc agg gga caa aag gga gac gta aac gtg tac aac 1104
Gln Lys Gln Ala Pro Arg Gly Gln Lys Gly Asp Val Asn Val Tyr Asn
355 360 365

tcc ttt ttc ctc caa cac gga ctg ggg ttt tta aat gag gcc acg atc 1152
Ser Phe Phe Leu Gln His Gly Leu Gly Phe Leu Asn Glu Ala Thr Ile
370 375 380

aag gaa aac ggc agc gaa gcc ttt aag ggc gta ccc tca aac gcc ctc 1200
Lys Glu Asn Gly Ser Glu Ala Phe Lys Gly Val Pro Ser Asn Ala Leu
385 390 395 400

gat ggt tct tcg ttc acg ccg tat cac ctg gcc tac gcc gcg tct ttc 1248
Asp Gly Ser Ser Phe Thr Pro Tyr His Leu Ala Tyr Ala Ala Ser Phe
405 410 415

tcg ccc cat ctg ctg gcg aag tta tgt tat tac atg cag ttc ttg caa 1296
Ser Pro His Leu Leu Ala Lys Leu Cys Tyr Tyr Met Gln Phe Leu Gln
420 425 430

cac cac aaa agc tcc acg aac cag gcg ttt aac atg gtc cat tat gtc 1344
His His Lys Ser Ser Thr Asn Gln Ala Phe Asn Met Val His Tyr Val
435 440 445

ggc acc gcc gcc aac tca gag atg tgc acg cta tgt cac ggc aac acg 1392
Gly Thr Ala Ala Asn Ser Glu Met Cys Thr Leu Cys His Gly Asn Thr
450 455 460

ccg gca acg tgc ctc aac acg ctg ttc tat aga ctg aag gat agg ttt 1440
Pro Ala Thr Cys Leu Asn Thr Leu Phe Tyr Arg Leu Lys Asp Arg Phe
465 470 475 480

ccc gcc gta acc acc cct cag cgc agg gac ccc tac gtg gtg acc gga 1488

Pro Ala Val Thr Thr Pro Gln Arg Arg Asp Pro Tyr Val Val Thr Gly
485 490 495

aca gcc ggg acc ttt aac gac ctg gag att ctg ggc aac ttc gcg agc 1536
Thr Ala Gly Thr Phe Asn Asp Leu Glu Ile Leu Gly Asn Phe Ala Ser
500 505 510

ttt aga gac cgc gaa gag gac gga aac ccg gcc gac gag cac cca aag 1584
Phe Arg Asp Arg Glu Glu Asp Gly Asn Pro Ala Asp Glu His Pro Lys
515 520 525

tac acg tac tgg cag cta tgt cag acc gtg aca gaa aag cta tcc gcg 1632
Tyr Thr Tyr Trp Gln Leu Cys Gln Thr Val Thr Glu Lys Leu Ser Ala
530 535 540

att gga atc acc gaa gac cac gat aat cac gtg aac ctc atc acc aac 1680
Ile Gly Ile Thr Glu Asp His Asp Asn His Val Asn Leu Ile Thr Asn
545 550 555 560

atc caa agt ttt ctc agg gtg ttc aag ggt atc gac tca att gtg gac 1728
Ile Gln Ser Phe Leu Arg Val Phe Lys Gly Ile Asp Ser Ile Val Asp
565 570 575

gga gag gtc atg aag ttc gtt aat tcg atg att aaa aat aac ttt aat 1776
Gly Glu Val Met Lys Phe Val Asn Ser Met Ile Lys Asn Asn Phe Asn
580 585 590

ttc cgc gag cac gtc aaa tcg gtc cat cac ata ctc cag ttc tgc tgc 1824
Phe Arg Glu His Val Lys Ser Val His His Ile Leu Gln Phe Cys Cys
595 600 605

aac gtg tat tgg cag gcg ccg tgc gcg gtg ttc ctg aat ctg tac tac 1872
Asn Val Tyr Trp Gln Ala Pro Cys Ala Val Phe Leu Asn Leu Tyr Tyr
610 615 620

aaa tcc ctg ctg tgg atc att cag gat atc tgc ctg ccg tac tgc atg 1920
Lys Ser Leu Leu Trp Ile Ile Gln Asp Ile Cys Leu Pro Tyr Cys Met
625 630 635 640

atc tac gaa caa gat aat ccg gcg atg ggc atc ctc ccc tcc gag tgg 1968
Ile Tyr Glu Gln Asp Asn Pro Ala Met Gly Ile Leu Pro Ser Glu Trp
645 650 655

cta aag atg cat ttt cag acg ttg tgg acg aac ttt aaa gcg gcg tgt 2016
Leu Lys Met His Phe Gln Thr Leu Trp Thr Asn Phe Lys Ala Ala Cys
660 665 670

ctc gac cgc ggg gtc ctc acg ggg tgc gaa ctg aaa atc gta cac cgg 2064
Leu Asp Arg Gly Val Leu Thr Gly Cys Glu Leu Lys Ile Val His Arg
675 680 685

gac atg ttc tgc gac ttc ttc gac acc gac gcg ggg tcc aac ggc tta 2112
Asp Met Phe Cys Asp Phe Phe Asp Thr Asp Ala Gly Ser Asn Gly Leu
690 695 700

atg gcc ccc ttt aaa atg cag gtt aga ata gcc cga gcc atg atg gtc 2160
Met Ala Pro Phe Lys Met Gln Val Arg Ile Ala Arg Ala Met Met Val
705 710 715 720

gtt ccg aaa tca att aaa ata aaa aat aga atc att ttt tcc aac acc 2208
Val Pro Lys Ser Ile Lys Ile Lys Asn Arg Ile Ile Phe Ser Asn Thr

```
              725                 730                 735

gcg gga tcc gag gcg gtg cag tcg ggg ttc gtc aaa ccg acg gga acc     2256
Ala Gly Ser Glu Ala Val Gln Ser Gly Phe Val Lys Pro Thr Gly Thr
            740                 745                 750

agg gac act tac gtg gtg gcc gga ccg tac atg aag ttt ctc aac tcg     2304
Arg Asp Thr Tyr Val Val Ala Gly Pro Tyr Met Lys Phe Leu Asn Ser
        755                 760                 765

ctg cat cgc gcg ctg ttc ccc gac acc aag acc gcc gcg ctg tac ctg     2352
Leu His Arg Ala Leu Phe Pro Asp Thr Lys Thr Ala Ala Leu Tyr Leu
    770                 775                 780

tgg cac aag atc tcc cag acc aac aaa acc cca gtt ctg aaa gac gtc     2400
Trp His Lys Ile Ser Gln Thr Asn Lys Thr Pro Val Leu Lys Asp Val
785                 790                 795                 800

ccg gac gac gag ctg gcg gag ctg gtg tcg tac gta aag acc aac agc     2448
Pro Asp Asp Glu Leu Ala Glu Leu Val Ser Tyr Val Lys Thr Asn Ser
                805                 810                 815

ctc gcg ttc gag gaa acg aac gtg ctg gac gtg gtt ccg gat tca ctc     2496
Leu Ala Phe Glu Glu Thr Asn Val Leu Asp Val Val Pro Asp Ser Leu
            820                 825                 830

atg tcg tac gcg agg atc aaa ctg aac ggg gcc att cta agg gca tgt     2544
Met Ser Tyr Ala Arg Ile Lys Leu Asn Gly Ala Ile Leu Arg Ala Cys
        835                 840                 845

ggc cag att cag ttc tac gcc acg acg ctg cac tgc ctc acg ccg gtg     2592
Gly Gln Ile Gln Phe Tyr Ala Thr Thr Leu His Cys Leu Thr Pro Val
    850                 855                 860

cta cag acg atc gat gcc gag gaa tac ccc cac gtg ctg ggc tcc gcg     2640
Leu Gln Thr Ile Asp Ala Glu Glu Tyr Pro His Val Leu Gly Ser Ala
865                 870                 875                 880

gca atc gcc aca ccg gtg gct tac ctg gca gaa ata cgc ggc cgc acc     2688
Ala Ile Ala Thr Pro Val Ala Tyr Leu Ala Glu Ile Arg Gly Arg Thr
                885                 890                 895

gcc ctc acc gtc cag acg acg gcg cgt cag ccg gtc gcc gcc aca ggg     2736
Ala Leu Thr Val Gln Thr Thr Ala Arg Gln Pro Val Ala Ala Thr Gly
            900                 905                 910

cgc ctg cgt ccc gtg ata acc gtt ccc atg gta gtc aac aaa tac acg     2784
Arg Leu Arg Pro Val Ile Thr Val Pro Met Val Val Asn Lys Tyr Thr
        915                 920                 925

ggg gtc aac ggg aac aac aac gtt ttc cac tgc gga aac ctg ggg tac     2832
Gly Val Asn Gly Asn Asn Asn Val Phe His Cys Gly Asn Leu Gly Tyr
    930                 935                 940

ttc gcg ggg cgc ggc gtg gac cgc aac ctg tgg ccg gaa agc tcc ccc     2880
Phe Ala Gly Arg Gly Val Asp Arg Asn Leu Trp Pro Glu Ser Ser Pro
945                 950                 955                 960

ttt aag aaa acg ggc gtc agc gcc atg cta aga aag aga cac gtc atg     2928
Phe Lys Lys Thr Gly Val Ser Ala Met Leu Arg Lys Arg His Val Met
                965                 970                 975
```

atg acc ccc att atc gac cgc cta ata aag cga gcc gcg gga cag aca 2976
Met Thr Pro Ile Ile Asp Arg Leu Ile Lys Arg Ala Ala Gly Gln Thr
980 985 990

atc agc acg ttc gag gcg gaa agc gtt aaa agg agc gtg cag gcg ctg 3024
Ile Ser Thr Phe Glu Ala Glu Ser Val Lys Arg Ser Val Gln Ala Leu
995 1000 1005

tta gag gat aag gac aac cct aac cta ttg aag tcg gta atc ttg gag 3072
Leu Glu Asp Lys Asp Asn Pro Asn Leu Leu Lys Ser Val Ile Leu Glu
1010 1015 1020

ctt ata cga cac ctg ggg aag ggc tgc cag gac tta agc tcc gag gac 3120
Leu Ile Arg His Leu Gly Lys Gly Cys Gln Asp Leu Ser Ser Glu Asp
1025 1030 1035 1040

gtg caa tat tac ctc ggt gac tat tgt atg ttg acg gac gag gtt tta 3168
Val Gln Tyr Tyr Leu Gly Asp Tyr Cys Met Leu Thr Asp Glu Val Leu
1045 1050 1055

ttt acg ttg gat aat ata gca cag tca ggc gtg ccg tgg act atc gag 3216
Phe Thr Leu Asp Asn Ile Ala Gln Ser Gly Val Pro Trp Thr Ile Glu
1060 1065 1070

gac gcg ggt gcc cta ata gag gat cgc cag gac gca gac gat ctt cag 3264
Asp Ala Gly Ala Leu Ile Glu Asp Arg Gln Asp Ala Asp Asp Leu Gln
1075 1080 1085

ttc gta gac agc gac gat atc gcc acc gct tcc tgt cag ccc ccc gag 3312
Phe Val Asp Ser Asp Asp Ile Ala Thr Ala Ser Cys Gln Pro Pro Glu
1090 1095 1100

gaa cag cta ccg acc cct agc gcc ggc gcc cta ctg gcc ggg aag aag 3360
Glu Gln Leu Pro Thr Pro Ser Ala Gly Ala Leu Leu Ala Gly Lys Lys
1105 1110 1115 1120

cga aaa att aac gcg ctg ctg agc gat cta gac ctt tag 3399
Arg Lys Ile Asn Ala Leu Leu Ser Asp Leu Asp Leu
1125 1130

<210> 9
<211> 1132
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 9
Met Ala Ser Lys Gly Asn Ala Gly Gln Pro Leu Glu Asp Asn Gln Gly
1 5 10 15
Ser Arg Ala Pro Ile Gly Ala Cys Gly Tyr Val Tyr Ala Tyr Ser Lys
20 25 30
Gln Asp Phe Pro Phe Ala Glu Ala Ser Ile Leu Gly Asn Arg Pro Ser
35 40 45
Gly Ser Gly Val Phe Ser Leu Pro Ile Leu Tyr Gly Leu Thr Val Glu
50 55 60
His Glu Phe Pro Leu Thr Val Lys Ala Ala Tyr Lys Lys Val Asp Thr
65 70 75 80
Thr Thr Leu Ala Val Lys Val Thr Cys Phe His Arg Glu Val Ile Val
85 90 95
Phe His Asn Ala Ser Leu Phe Arg Pro Val Phe Asp Gly Thr Gly Leu

100 105 110
Asn Glu Leu Cys Glu Glu Ala Arg Ala Leu Phe Gly Tyr Thr Gln Phe
115 120 125
Ile Glu Pro Gly Pro Pro His Ser Ile Trp Asn Pro Leu Glu Cys Pro
130 135 140
Gln Leu Pro Asp Lys Asp Glu Met Phe Leu Gly Val Val Val Thr Glu
145 150 155 160
Gly Phe Lys Glu Arg Leu Trp Arg Gly Cys Leu Val Pro Ala Val Phe
165 170 175
Gln Thr Gln Gln Val Gln Ile Ala Gly Arg Gln Ala Phe Lys Val Pro
180 185 190
Leu Tyr Asp Glu Asp Leu Phe Ala Pro His Gly His Arg Met Pro Arg
195 200 205
Phe Tyr His Lys Asp Val Ser Ala Tyr Leu Tyr Asp Ser Leu Phe Thr
210 215 220
Ser Ile Ala Gln Ala Leu Arg Leu Lys Asp Val Thr Ala Val Ile His
225 230 235 240
Ala Thr Glu Lys Gln Phe Met Gln Asp His Tyr Lys Ile Ala Lys Ile
245 250 255
Val Gln Ala Lys Gln Phe Ser Thr Thr Leu Pro Lys Thr Thr Asp Gly
260 265 270
Ser Ser His Met Ile Val Asp Ser Val Val Ala Glu Leu Ala Leu Ser
275 280 285
Tyr Gly Cys Met Phe Leu Glu Cys Pro Gln Asp Ala Cys Glu Leu Leu
290 295 300
Asn Tyr Asp Ser Trp Pro Ile Phe Asp Gly Cys Asp Ser Pro Glu Ala
305 310 315 320
Arg Val Asn Ala Leu Glu Arg Trp Ser Ala Glu Gln Ala Val His Val
325 330 335
Ala Gly Gln Leu Phe Ala Ala Asn Ser Val Leu Tyr Leu Thr Lys Val
340 345 350
Gln Lys Gln Ala Pro Arg Gly Gln Lys Gly Asp Val Asn Val Tyr Asn
355 360 365
Ser Phe Phe Leu Gln His Gly Leu Gly Phe Leu Asn Glu Ala Thr Ile
370 375 380
Lys Glu Asn Gly Ser Glu Ala Phe Lys Gly Val Pro Ser Asn Ala Leu
385 390 395 400
Asp Gly Ser Ser Phe Thr Pro Tyr His Leu Ala Tyr Ala Ala Ser Phe
405 410 415
Ser Pro His Leu Leu Ala Lys Leu Cys Tyr Tyr Met Gln Phe Leu Gln
420 425 430
His His Lys Ser Ser Thr Asn Gln Ala Phe Asn Met Val His Tyr Val
435 440 445
Gly Thr Ala Ala Asn Ser Glu Met Cys Thr Leu Cys His Gly Asn Thr
450 455 460
Pro Ala Thr Cys Leu Asn Thr Leu Phe Tyr Arg Leu Lys Asp Arg Phe
465 470 475 480
Pro Ala Val Thr Thr Pro Gln Arg Arg Asp Pro Tyr Val Val Thr Gly
485 490 495
Thr Ala Gly Thr Phe Asn Asp Leu Glu Ile Leu Gly Asn Phe Ala Ser
500 505 510
Phe Arg Asp Arg Glu Glu Asp Gly Asn Pro Ala Asp Glu His Pro Lys
515 520 525
Tyr Thr Tyr Trp Gln Leu Cys Gln Thr Val Thr Glu Lys Leu Ser Ala
530 535 540
Ile Gly Ile Thr Glu Asp His Asp Asn His Val Asn Leu Ile Thr Asn
545 550 555 560
Ile Gln Ser Phe Leu Arg Val Phe Lys Gly Ile Asp Ser Ile Val Asp
565 570 575
Gly Glu Val Met Lys Phe Val Asn Ser Met Ile Lys Asn Asn Phe Asn
580 585 590

Phe Arg Glu His Val Lys Ser Val His His Ile Leu Gln Phe Cys Cys
        595   600   605
Asn Val Tyr Trp Gln Ala Pro Cys Ala Val Phe Leu Asn Leu Tyr Tyr
    610   615   620
Lys Ser Leu Leu Trp Ile Ile Gln Asp Ile Cys Leu Pro Tyr Cys Met
625   630   635   640
Ile Tyr Glu Gln Asp Asn Pro Ala Met Gly Ile Leu Pro Ser Glu Trp
                645   650   655
Leu Lys Met His Phe Gln Thr Leu Trp Thr Asn Phe Lys Ala Ala Cys
            660   665   670
Leu Asp Arg Gly Val Leu Thr Gly Cys Glu Leu Lys Ile Val His Arg
        675   680   685
Asp Met Phe Cys Asp Phe Phe Asp Thr Asp Ala Gly Ser Asn Gly Leu
    690   695   700
Met Ala Pro Phe Lys Met Gln Val Arg Ile Ala Arg Ala Met Met Val
705   710   715   720
Val Pro Lys Ser Ile Lys Ile Lys Asn Arg Ile Ile Phe Ser Asn Thr
                725   730   735
Ala Gly Ser Glu Ala Val Gln Ser Gly Phe Val Lys Pro Thr Gly Thr
            740   745   750
Arg Asp Thr Tyr Val Val Ala Gly Pro Tyr Met Lys Phe Leu Asn Ser
        755   760   765
Leu His Arg Ala Leu Phe Pro Asp Thr Lys Thr Ala Ala Leu Tyr Leu
    770   775   780
Trp His Lys Ile Ser Gln Thr Asn Lys Thr Pro Val Leu Lys Asp Val
785   790   795   800
Pro Asp Asp Glu Leu Ala Glu Leu Val Ser Tyr Val Lys Thr Asn Ser
                805   810   815
Leu Ala Phe Glu Glu Thr Asn Val Leu Asp Val Val Pro Asp Ser Leu
            820   825   830
Met Ser Tyr Ala Arg Ile Lys Leu Asn Gly Ala Ile Leu Arg Ala Cys
        835   840   845
Gly Gln Ile Gln Phe Tyr Ala Thr Thr Leu His Cys Leu Thr Pro Val
    850   855   860
Leu Gln Thr Ile Asp Ala Glu Glu Tyr Pro His Val Leu Gly Ser Ala
865   870   875   880
Ala Ile Ala Thr Pro Val Ala Tyr Leu Ala Glu Ile Arg Gly Arg Thr
                885   890   895
Ala Leu Thr Val Gln Thr Thr Ala Arg Gln Pro Val Ala Ala Thr Gly
            900   905   910
Arg Leu Arg Pro Val Ile Thr Val Pro Met Val Val Asn Lys Tyr Thr
        915   920   925
Gly Val Asn Gly Asn Asn Asn Val Phe His Cys Gly Asn Leu Gly Tyr
    930   935   940
Phe Ala Gly Arg Gly Val Asp Arg Asn Leu Trp Pro Glu Ser Ser Pro
945   950   955   960
Phe Lys Lys Thr Gly Val Ser Ala Met Leu Arg Lys Arg His Val Met
                965   970   975
Met Thr Pro Ile Ile Asp Arg Leu Ile Lys Arg Ala Ala Gly Gln Thr
            980   985   990
Ile Ser Thr Phe Glu Ala Glu Ser Val Lys Arg Ser Val Gln Ala Leu
        995   1000   1005
Leu Glu Asp Lys Asp Asn Pro Asn Leu Leu Lys Ser Val Ile Leu Glu
   1010   1015   1020
Leu Ile Arg His Leu Gly Lys Gly Cys Gln Asp Leu Ser Ser Glu Asp
1025   1030   1035   1040
Val Gln Tyr Tyr Leu Gly Asp Tyr Cys Met Leu Thr Asp Glu Val Leu
               1045   1050   1055
Phe Thr Leu Asp Asn Ile Ala Gln Ser Gly Val Pro Trp Thr Ile Glu
           1060   1065   1070
Asp Ala Gly Ala Leu Ile Glu Asp Arg Gln Asp Ala Asp Asp Leu Gln

1075 1080 1085
Phe Val Asp Ser Asp Asp Ile Ala Thr Ala Ser Cys Gln Pro Pro Glu
1090 1095 1100
Glu Gln Leu Pro Thr Pro Ser Ala Gly Ala Leu Leu Ala Gly Lys Lys
1105 1110 1115 1120
Arg Lys Ile Asn Ala Leu Leu Ser Asp Leu Asp Leu
1125 1130

<210> 10
<211> 2061
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2061)

<400> 10
atg gcc agg gaa ctc gca gca tta tac gcg cag ctg tcg gcc ctc gcc 48
Met Ala Arg Glu Leu Ala Ala Leu Tyr Ala Gln Leu Ser Ala Leu Ala
1 5 10 15

gtc gac ttg agt ctg gtt atc ttt gcg gac ccg cga agt atc gac ggt 96
Val Asp Leu Ser Leu Val Ile Phe Ala Asp Pro Arg Ser Ile Asp Gly
20 25 30

gcc cgc att cta aaa aca aaa aca cag ata gag aac ctg aac cgc gac 144
Ala Arg Ile Leu Lys Thr Lys Thr Gln Ile Glu Asn Leu Asn Arg Asp
35 40 45

ctt ctg ccg ctg cta cgc gag caa aac tcg gta gag acg tcc agc ctg 192
Leu Leu Pro Leu Leu Arg Glu Gln Asn Ser Val Glu Thr Ser Ser Leu
50 55 60

tcg ctc gaa gtg gag cac ctg gcc aaa aac atc gag gac aaa ctc ggc 240
Ser Leu Glu Val Glu His Leu Ala Lys Asn Ile Glu Asp Lys Leu Gly
65 70 75 80

gag ctg gag cgc agt ctg cgg cag aga tat tcg agc cga gag cat ttt 288
Glu Leu Glu Arg Ser Leu Arg Gln Arg Tyr Ser Ser Arg Glu His Phe
85 90 95

gaa aca cta cac ctg aga ccc gaa tgt cac tat cac tct acg gtt act 336
Glu Thr Leu His Leu Arg Pro Glu Cys His Tyr His Ser Thr Val Thr
100 105 110

ttt cag ttt tac ggg ggc ggg tta ata gat gta aac atg tgc cta ata 384
Phe Gln Phe Tyr Gly Gly Gly Leu Ile Asp Val Asn Met Cys Leu Ile
115 120 125

aac gat gta gaa ctg ctg tgt aaa aga cta ggg agt gtg ttt tat tgc 432
Asn Asp Val Glu Leu Leu Cys Lys Arg Leu Gly Ser Val Phe Tyr Cys
130 135 140

atc ggt gca aac gaa gct ctg tcc gga ttg aac cgg gtt ctg acg ttt 480
Ile Gly Ala Asn Glu Ala Leu Ser Gly Leu Asn Arg Val Leu Thr Phe
145 150 155 160

ctg tca aca ctg cgg ggt atc tcc ccg atc ccg cac cca gac cta tac 528

Leu Ser Thr Leu Arg Gly Ile Ser Pro Ile Pro His Pro Asp Leu Tyr
165 170 175

gtc acg tca gtg cct tgc gta cag tgc ctg agg gaa atc gaa ctc gta 576
Val Thr Ser Val Pro Cys Val Gln Cys Leu Arg Glu Ile Glu Leu Val
180 185 190

cca aat cag ggg tcc agt tta ctc gcg gtg ttg gca gac cga cac tgc 624
Pro Asn Gln Gly Ser Ser Leu Leu Ala Val Leu Ala Asp Arg His Cys
195 200 205

gat cac ctc tgt aag aag gtt agg gcg gag cca ata cac ggc ctg ttt 672
Asp His Leu Cys Lys Lys Val Arg Ala Glu Pro Ile His Gly Leu Phe
210 215 220

gag aca gaa ctg agc cag ctg ggt cta aaa gta aca aaa cgt tcg gac 720
Glu Thr Glu Leu Ser Gln Leu Gly Leu Lys Val Thr Lys Arg Ser Asp
225 230 235 240

gcc acg cag cac ggc gtc cgg tcc tct gca gat cag tta agg gag tcg 768
Ala Thr Gln His Gly Val Arg Ser Ser Ala Asp Gln Leu Arg Glu Ser
245 250 255

tcg ctg gcg gcc ata caa gat cac aat ata ttc aaa cgg gtg tcc gcg 816
Ser Leu Ala Ala Ile Gln Asp His Asn Ile Phe Lys Arg Val Ser Ala
260 265 270

tca atc atg gaa cta tcc aat cta att tat tgg aac gcc ggg caa acc 864
Ser Ile Met Glu Leu Ser Asn Leu Ile Tyr Trp Asn Ala Gly Gln Thr
275 280 285

ggc ctc cag acc ggg acc gaa aac gag tgc tca caa atg gcc aga ctg 912
Gly Leu Gln Thr Gly Thr Glu Asn Glu Cys Ser Gln Met Ala Arg Leu
290 295 300

cta aca cac gag gcc gat atg cac gag cac cgt gcg cta ata aca ccc 960
Leu Thr His Glu Ala Asp Met His Glu His Arg Ala Leu Ile Thr Pro
305 310 315 320

aaa cta agc gcg act cac ttc tac gac tgt ttc cga ccg gat ccc ata 1008
Lys Leu Ser Ala Thr His Phe Tyr Asp Cys Phe Arg Pro Asp Pro Ile
325 330 335

gaa tcc ctg ttc tgc ggc ggt ctt ttt aac tct ata gac gac acc ata 1056
Glu Ser Leu Phe Cys Gly Gly Leu Phe Asn Ser Ile Asp Asp Thr Ile
340 345 350

aac gca ctg agc cgg gat tgc tcc gtg acg ttc ttt caa cag gca aac 1104
Asn Ala Leu Ser Arg Asp Cys Ser Val Thr Phe Phe Gln Gln Ala Asn
355 360 365

tat acc aac gtt atg cga aaa caa aac gag ctg ttc acc aga ctc aat 1152
Tyr Thr Asn Val Met Arg Lys Gln Asn Glu Leu Phe Thr Arg Leu Asn
370 375 380

agc atc ctg cgt cag ggg agc gcg gga tcg caa aaa ccg gcc acc ccc 1200
Ser Ile Leu Arg Gln Gly Ser Ala Gly Ser Gln Lys Pro Ala Thr Pro
385 390 395 400

tcg gag cca cgg acc acc acc gtg gcg gca acc gcg gca agc gac gtc 1248
Ser Glu Pro Arg Thr Thr Thr Val Ala Ala Thr Ala Ala Ser Asp Val

```
            405                 410                 415

atc aaa gac gca cag tat cgc aaa gaa cag tac atg aaa aag gtg gcc    1296
Ile Lys Asp Ala Gln Tyr Arg Lys Glu Gln Tyr Met Lys Lys Val Ala
        420                 425                 430

agg gac ggc ttt aaa aaa cta aca gag tgt ctg cag acg cag agc gcg    1344
Arg Asp Gly Phe Lys Lys Leu Thr Glu Cys Leu Gln Thr Gln Ser Ala
        435                 440                 445

gtg ttg gca aac gca ctc tgc atg cgc gta tgg ggg ggc gtc gcc tac    1392
Val Leu Ala Asn Ala Leu Cys Met Arg Val Trp Gly Gly Val Ala Tyr
    450                 455                 460

ggc gag gcg tcc gag ctg gtg aac cat ttt ctc ctc agg cgg cgc ttc    1440
Gly Glu Ala Ser Glu Leu Val Asn His Phe Leu Leu Arg Arg Arg Phe
465                 470                 475                 480

gtc gcg ctt ccc tgg gag gcg cgc tgc cgc tcg gat cag att tta ttc    1488
Val Ala Leu Pro Trp Glu Ala Arg Cys Arg Ser Asp Gln Ile Leu Phe
            485                 490                 495

gaa aac tca aag tac att aaa aac tca cta tat tcc cag cgc ctc agt    1536
Glu Asn Ser Lys Tyr Ile Lys Asn Ser Leu Tyr Ser Gln Arg Leu Ser
        500                 505                 510

cgc gaa cac gta gag att atc acg ctg cag ttt tac ggc ctg ata acc    1584
Arg Glu His Val Glu Ile Ile Thr Leu Gln Phe Tyr Gly Leu Ile Thr
        515                 520                 525

ggc ccc ctg acg cgc cag agc gat ctc ttt ccc ggc ccc gcc aac gtc    1632
Gly Pro Leu Thr Arg Gln Ser Asp Leu Phe Pro Gly Pro Ala Asn Val
    530                 535                 540

gcg ctg gcc cag tgt ttc gag gcg gcc gga atg ctt ccg cat cac aag    1680
Ala Leu Ala Gln Cys Phe Glu Ala Ala Gly Met Leu Pro His His Lys
545                 550                 555                 560

atg ctg gtg tca gag atg ata tgg ccc cag att caa ccg aaa gac tgg    1728
Met Leu Val Ser Glu Met Ile Trp Pro Gln Ile Gln Pro Lys Asp Trp
            565                 570                 575

ata gac cag aca ttt aat cgt ttt tac caa ctt ccc gaa ggt gat ctc    1776
Ile Asp Gln Thr Phe Asn Arg Phe Tyr Gln Leu Pro Glu Gly Asp Leu
        580                 585                 590

aac gcg gta caa aag tcc gcc tgg tgc ttt ata cga gag ctc gtc ctc    1824
Asn Ala Val Gln Lys Ser Ala Trp Cys Phe Ile Arg Glu Leu Val Leu
        595                 600                 605

tcg gtg gcg ctt tat aat cgc acg tgg gaa aag acg ctg cgg ata ttt    1872
Ser Val Ala Leu Tyr Asn Arg Thr Trp Glu Lys Thr Leu Arg Ile Phe
    610                 615                 620

tcc cta gcg cgc gag aaa ctc tcc atc tcc aac cta gac gtt aaa ggc    1920
Ser Leu Ala Arg Glu Lys Leu Ser Ile Ser Asn Leu Asp Val Lys Gly
625                 630                 635                 640

ctg acg tcc ggc ctg tat cta acg tac gag caa gac gcg ccg ctc gtt    1968
Leu Thr Ser Gly Leu Tyr Leu Thr Tyr Glu Gln Asp Ala Pro Leu Val
            645                 650                 655
```

```
cta att tct caa aat acc ggc tgg ata ttt aaa gac ctg tac gct ctt     2016
Leu Ile Ser Gln Asn Thr Gly Trp Ile Phe Lys Asp Leu Tyr Ala Leu
            660                 665                 670

ctg tac cat cac ctg caa ctg tcc gac ggc cat gat gat aac taa         2061
Leu Tyr His His Leu Gln Leu Ser Asp Gly His Asp Asp Asn
        675                 680                 685


<210> 11
<211> 686
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 11
Met Ala Arg Glu Leu Ala Ala Leu Tyr Ala Gln Leu Ser Ala Leu Ala
  1               5                  10                  15
Val Asp Leu Ser Leu Val Ile Phe Ala Asp Pro Arg Ser Ile Asp Gly
            20                  25                  30
Ala Arg Ile Leu Lys Thr Lys Thr Gln Ile Glu Asn Leu Asn Arg Asp
        35                  40                  45
Leu Leu Pro Leu Leu Arg Glu Gln Asn Ser Val Glu Thr Ser Ser Leu
    50                  55                  60
Ser Leu Glu Val Glu His Leu Ala Lys Asn Ile Glu Asp Lys Leu Gly
 65                  70                  75                  80
Glu Leu Glu Arg Ser Leu Arg Gln Arg Tyr Ser Ser Arg Glu His Phe
                 85                  90                  95
Glu Thr Leu His Leu Arg Pro Glu Cys His Tyr His Ser Thr Val Thr
            100                 105                 110
Phe Gln Phe Tyr Gly Gly Gly Leu Ile Asp Val Asn Met Cys Leu Ile
        115                 120                 125
Asn Asp Val Glu Leu Leu Cys Lys Arg Leu Gly Ser Val Phe Tyr Cys
    130                 135                 140
Ile Gly Ala Asn Glu Ala Leu Ser Gly Leu Asn Arg Val Leu Thr Phe
145                 150                 155                 160
Leu Ser Thr Leu Arg Gly Ile Ser Pro Ile Pro His Pro Asp Leu Tyr
                165                 170                 175
Val Thr Ser Val Pro Cys Val Gln Cys Leu Arg Glu Ile Glu Leu Val
            180                 185                 190
Pro Asn Gln Gly Ser Ser Leu Leu Ala Val Leu Ala Asp Arg His Cys
        195                 200                 205
Asp His Leu Cys Lys Lys Val Arg Ala Glu Pro Ile His Gly Leu Phe
    210                 215                 220
Glu Thr Glu Leu Ser Gln Leu Gly Leu Lys Val Thr Lys Arg Ser Asp
225                 230                 235                 240
Ala Thr Gln His Gly Val Arg Ser Ser Ala Asp Gln Leu Arg Glu Ser
                245                 250                 255
Ser Leu Ala Ala Ile Gln Asp His Asn Ile Phe Lys Arg Val Ser Ala
            260                 265                 270
Ser Ile Met Glu Leu Ser Asn Leu Ile Tyr Trp Asn Ala Gly Gln Thr
        275                 280                 285
Gly Leu Gln Thr Gly Thr Glu Asn Glu Cys Ser Gln Met Ala Arg Leu
    290                 295                 300
Leu Thr His Glu Ala Asp Met His Glu His Arg Ala Leu Ile Thr Pro
305                 310                 315                 320
Lys Leu Ser Ala Thr His Phe Tyr Asp Cys Phe Arg Pro Asp Pro Ile
                325                 330                 335
Glu Ser Leu Phe Cys Gly Gly Leu Phe Asn Ser Ile Asp Asp Thr Ile
            340                 345                 350
Asn Ala Leu Ser Arg Asp Cys Ser Val Thr Phe Phe Gln Gln Ala Asn
```

```
                355                 360                 365
Tyr Thr Asn Val Met Arg Lys Gln Asn Glu Leu Phe Thr Arg Leu Asn
            370                 375                 380
Ser Ile Leu Arg Gln Gly Ser Ala Gly Ser Gln Lys Pro Ala Thr Pro
385                 390                 395                 400
Ser Glu Pro Arg Thr Thr Thr Val Ala Ala Thr Ala Ala Ser Asp Val
                405                 410                 415
Ile Lys Asp Ala Gln Tyr Arg Lys Glu Gln Tyr Met Lys Lys Val Ala
            420                 425                 430
Arg Asp Gly Phe Lys Lys Leu Thr Glu Cys Leu Gln Thr Gln Ser Ala
        435                 440                 445
Val Leu Ala Asn Ala Leu Cys Met Arg Val Trp Gly Gly Val Ala Tyr
    450                 455                 460
Gly Glu Ala Ser Glu Leu Val Asn His Phe Leu Leu Arg Arg Arg Phe
465                 470                 475                 480
Val Ala Leu Pro Trp Glu Ala Arg Cys Arg Ser Asp Gln Ile Leu Phe
                485                 490                 495
Glu Asn Ser Lys Tyr Ile Lys Asn Ser Leu Tyr Ser Gln Arg Leu Ser
            500                 505                 510
Arg Glu His Val Glu Ile Ile Thr Leu Gln Phe Tyr Gly Leu Ile Thr
        515                 520                 525
Gly Pro Leu Thr Arg Gln Ser Asp Leu Phe Pro Gly Pro Ala Asn Val
    530                 535                 540
Ala Leu Ala Gln Cys Phe Glu Ala Ala Gly Met Leu Pro His His Lys
545                 550                 555                 560
Met Leu Val Ser Glu Met Ile Trp Pro Gln Ile Gln Pro Lys Asp Trp
                565                 570                 575
Ile Asp Gln Thr Phe Asn Arg Phe Tyr Gln Leu Pro Glu Gly Asp Leu
            580                 585                 590
Asn Ala Val Gln Lys Ser Ala Trp Cys Phe Ile Arg Glu Leu Val Leu
        595                 600                 605
Ser Val Ala Leu Tyr Asn Arg Thr Trp Glu Lys Thr Leu Arg Ile Phe
    610                 615                 620
Ser Leu Ala Arg Glu Lys Leu Ser Ile Ser Asn Leu Asp Val Lys Gly
625                 630                 635                 640
Leu Thr Ser Gly Leu Tyr Leu Thr Tyr Glu Gln Asp Ala Pro Leu Val
                645                 650                 655
Leu Ile Ser Gln Asn Thr Gly Trp Ile Phe Lys Asp Leu Tyr Ala Leu
            660                 665                 670
Leu Tyr His His Leu Gln Leu Ser Asp Gly His Asp Asp Asn
        675                 680                 685


<210> 12
<211> 2490
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2490)

<400> 12
atg atg ata act aac cga acg cgt cgt ctc ctg cgg gcg tgg gtc gtg      48
Met Met Ile Thr Asn Arg Thr Arg Arg Leu Leu Arg Ala Trp Val Val
  1               5                  10                  15

ata atc gcg atc ggc acg gcg gtt ggc gaa aac gtc acc acc ccc aag      96
Ile Ile Ala Ile Gly Thr Ala Val Gly Glu Asn Val Thr Thr Pro Lys
             20                  25                  30
```

```
ggc gcg acc acc acc gcg aag cca acg ccg ggc ccg tcg acg ccc aca    144
Gly Ala Thr Thr Thr Ala Lys Pro Thr Pro Gly Pro Ser Thr Pro Thr
         35                  40                  45

cct ccc gag aac cca cct agg gcc gag gcg ttt aag ttt cgc gtg tgc    192
Pro Pro Glu Asn Pro Pro Arg Ala Glu Ala Phe Lys Phe Arg Val Cys
     50                  55                  60

agc gcc tcg gcc acc ggc gaa ctc ttc agg ttt aac ctg gaa aaa acg    240
Ser Ala Ser Ala Thr Gly Glu Leu Phe Arg Phe Asn Leu Glu Lys Thr
 65                  70                  75                  80

tgt ccg ggc acc gag gac aag acg cac caa gaa ggc atc ctg atg gtg    288
Cys Pro Gly Thr Glu Asp Lys Thr His Gln Glu Gly Ile Leu Met Val
                 85                  90                  95

ttt aaa aaa aat att gtc ccg cac atc ttt aag gtc aga cgg tac cgc    336
Phe Lys Lys Asn Ile Val Pro His Ile Phe Lys Val Arg Arg Tyr Arg
            100                 105                 110

aag gtg gcc acc tcg gtg acc gtc tat cga ggg tgg acc gag acc gcc    384
Lys Val Ala Thr Ser Val Thr Val Tyr Arg Gly Trp Thr Glu Thr Ala
        115                 120                 125

gtg acc ggc aag caa gag gtc atc cga ccg gtg ccg cag tac gag atc    432
Val Thr Gly Lys Gln Glu Val Ile Arg Pro Val Pro Gln Tyr Glu Ile
    130                 135                 140

aac cac atg gac acg acc tac cag tgt ttc agc tcc atg cgc gta aac    480
Asn His Met Asp Thr Thr Tyr Gln Cys Phe Ser Ser Met Arg Val Asn
145                 150                 155                 160

gtc aac ggc ata gta aac acc tac acg gac agg gac ttc act aac cag    528
Val Asn Gly Ile Val Asn Thr Tyr Thr Asp Arg Asp Phe Thr Asn Gln
                165                 170                 175

acc gtg ttt ctg caa ccg gtc gag ggg ctc acg gat aac atc cag cga    576
Thr Val Phe Leu Gln Pro Val Glu Gly Leu Thr Asp Asn Ile Gln Arg
            180                 185                 190

tac ttc agt cag ccg gtg ctg tac acg aca ccg gga tgg ttt ccg gga    624
Tyr Phe Ser Gln Pro Val Leu Tyr Thr Thr Pro Gly Trp Phe Pro Gly
        195                 200                 205

att tac agg gtc aga acc acg gtc aac tgc gag atc gtg gac atg atc    672
Ile Tyr Arg Val Arg Thr Thr Val Asn Cys Glu Ile Val Asp Met Ile
    210                 215                 220

gcg cgt tcg gcg gaa ccg tac tcg tat ttt gtc acc gcc ctg gga gac    720
Ala Arg Ser Ala Glu Pro Tyr Ser Tyr Phe Val Thr Ala Leu Gly Asp
225                 230                 235                 240

acg gta gag gtg tcc ccg ttc tgc cac aac gac tca acg tgc tcg gtc    768
Thr Val Glu Val Ser Pro Phe Cys His Asn Asp Ser Thr Cys Ser Val
                245                 250                 255

gcg gag aaa acc gaa aac ggc ctc ggc gcc cgc gtg ctc aca aat tac    816
Ala Glu Lys Thr Glu Asn Gly Leu Gly Ala Arg Val Leu Thr Asn Tyr
            260                 265                 270
```

```
acc atc gtc gac ttc gcg acc cgc cag ccc acc acc gaa acg cgg gtc     864
Thr Ile Val Asp Phe Ala Thr Arg Gln Pro Thr Thr Glu Thr Arg Val
        275                 280                 285

ttc gcc gac tcg gga gaa tac acc gta tcg tgg aag gcg gag gac ccc     912
Phe Ala Asp Ser Gly Glu Tyr Thr Val Ser Trp Lys Ala Glu Asp Pro
    290                 295                 300

aag tcg gcg gtc tgc gcg ctg acg ctc tgg aaa acc ttc ccc agg gcg     960
Lys Ser Ala Val Cys Ala Leu Thr Leu Trp Lys Thr Phe Pro Arg Ala
305                 310                 315                 320

ata cag acg acg cac gag gcc agc tac cac ttc gtg gcc aac gac gtg     1008
Ile Gln Thr Thr His Glu Ala Ser Tyr His Phe Val Ala Asn Asp Val
                325                 330                 335

acg gcg acc ttc acg tcc ccg ctc tcc cag gta act aac ttc acg ggc     1056
Thr Ala Thr Phe Thr Ser Pro Leu Ser Gln Val Thr Asn Phe Thr Gly
            340                 345                 350

acg tac ccc tgc ctc aat gat gtt att cag aaa acc ctc aac gcc acc     1104
Thr Tyr Pro Cys Leu Asn Asp Val Ile Gln Lys Thr Leu Asn Ala Thr
        355                 360                 365

atc aag aag ctg tcc gat acc cac gca aca aac gga tcg gag cag tac     1152
Ile Lys Lys Leu Ser Asp Thr His Ala Thr Asn Gly Ser Glu Gln Tyr
    370                 375                 380

tac gaa acc gag ggg ggt ctg ttt ctc ctg tgg cag ccg tta acg ccg     1200
Tyr Glu Thr Glu Gly Gly Leu Phe Leu Leu Trp Gln Pro Leu Thr Pro
385                 390                 395                 400

cta agc cta gct gac gag atg cgc gaa tta aac ggc acc acg cca gca     1248
Leu Ser Leu Ala Asp Glu Met Arg Glu Leu Asn Gly Thr Thr Pro Ala
                405                 410                 415

ccc ccc acc aca acc tca acc gcc aac cgc gtt cga aga agc gtc ggt     1296
Pro Pro Thr Thr Thr Ser Thr Ala Asn Arg Val Arg Arg Ser Val Gly
            420                 425                 430

acg aac gag cag gca acg gac gac cta gcg gcg ccc cag ctg cag ttc     1344
Thr Asn Glu Gln Ala Thr Asp Asp Leu Ala Ala Pro Gln Leu Gln Phe
        435                 440                 445

gcc tac gac aag ctc cgc gcg agc atc aac aag gtg ctg gag gag ctc     1392
Ala Tyr Asp Lys Leu Arg Ala Ser Ile Asn Lys Val Leu Glu Glu Leu
    450                 455                 460

tcc agg gcg tgg tgc cga gaa cag gtg agg gac acc tac atg tgg tac     1440
Ser Arg Ala Trp Cys Arg Glu Gln Val Arg Asp Thr Tyr Met Trp Tyr
465                 470                 475                 480

gaa ctg agc aag att aac ccc acc agc gta atg acg gcg ata tac ggg     1488
Glu Leu Ser Lys Ile Asn Pro Thr Ser Val Met Thr Ala Ile Tyr Gly
                485                 490                 495

cgg ccg gtg tcg gcc aag ttc gtg ggc gac gcc atc tcc gtg acg gac     1536
Arg Pro Val Ser Ala Lys Phe Val Gly Asp Ala Ile Ser Val Thr Asp
            500                 505                 510

tgc gtg gcg gtg gac cag gcg tcc gtc agc atc cac aag agc ctc cgc     1584
```

Cys Val Ala Val Asp Gln Ala Ser Val Ser Ile His Lys Ser Leu Arg
515 520 525

acg tcc acc ccg ggg atc tgc tac tcg cgc ccc ccg gtc acg ttc agg 1632
Thr Ser Thr Pro Gly Ile Cys Tyr Ser Arg Pro Pro Val Thr Phe Arg
530 535 540

ttc ctc aac agc acc acg ctg ttc aag ggc cag ctg gga ccc aga aac 1680
Phe Leu Asn Ser Thr Thr Leu Phe Lys Gly Gln Leu Gly Pro Arg Asn
545 550 555 560

gag atc ata ctg acg gac aac cag gtg gag gcg tgc aaa gag acg tgc 1728
Glu Ile Ile Leu Thr Asp Asn Gln Val Glu Ala Cys Lys Glu Thr Cys
565 570 575

gaa cac tac ttc ata gcg agc aac gta acc tac tac tac aaa gac tac 1776
Glu His Tyr Phe Ile Ala Ser Asn Val Thr Tyr Tyr Tyr Lys Asp Tyr
580 585 590

gtc ttc gtg aaa aaa att aac acc tcc gag ata tcc acc ctc ggt acg 1824
Val Phe Val Lys Lys Ile Asn Thr Ser Glu Ile Ser Thr Leu Gly Thr
595 600 605

ttc atc gcc ctg aac ctg tcg ttt ata gag aac ata gat ttc agg gtc 1872
Phe Ile Ala Leu Asn Leu Ser Phe Ile Glu Asn Ile Asp Phe Arg Val
610 615 620

atc gag ctg tac agc cgc gcg gag aaa aag ctg tcc ggg agc gtt ttc 1920
Ile Glu Leu Tyr Ser Arg Ala Glu Lys Lys Leu Ser Gly Ser Val Phe
625 630 635 640

gat ata gaa acc atg ttc agg gaa tac aac tac tac acg caa cgc ctg 1968
Asp Ile Glu Thr Met Phe Arg Glu Tyr Asn Tyr Tyr Thr Gln Arg Leu
645 650 655

gcg gga ctc cgg gag gac ctg gac aac acg atc gac ctg aac cgc gac 2016
Ala Gly Leu Arg Glu Asp Leu Asp Asn Thr Ile Asp Leu Asn Arg Asp
660 665 670

cgc ctg gcc cgc gac ctg tcc gag ata gtc gcg gac ctg ggc gat gtc 2064
Arg Leu Ala Arg Asp Leu Ser Glu Ile Val Ala Asp Leu Gly Asp Val
675 680 685

ggc cgc acg gtc gtt aac gtg gcc agt agc gtg ata acc ctg ttc gga 2112
Gly Arg Thr Val Val Asn Val Ala Ser Ser Val Ile Thr Leu Phe Gly
690 695 700

tca atc gtg agc ggg ttc att aac ttt ata aag agt ccg ttc ggg ggc 2160
Ser Ile Val Ser Gly Phe Ile Asn Phe Ile Lys Ser Pro Phe Gly Gly
705 710 715 720

atg ctc atg atc ctg gtg att gtg gcg gtc gtc ctg atc gtg ttt gcg 2208
Met Leu Met Ile Leu Val Ile Val Ala Val Val Leu Ile Val Phe Ala
725 730 735

cta aac cgg cgc acc aac gcc atc gcc cag gcc ccc atc agg atg atc 2256
Leu Asn Arg Arg Thr Asn Ala Ile Ala Gln Ala Pro Ile Arg Met Ile
740 745 750

tac ccc gac ata gac aaa atg cag ccc tct ggc ggt aaa gtc gac cag 2304
Tyr Pro Asp Ile Asp Lys Met Gln Pro Ser Gly Gly Lys Val Asp Gln

```
        755                 760                 765

gag cag att aaa aac att ctc gcc ggc atg cac cag cta cag cag gaa      2352
Glu Gln Ile Lys Asn Ile Leu Ala Gly Met His Gln Leu Gln Gln Glu
    770                 775                 780

gag cgt agg cgg tta gac gaa cag cag agg tca gcg ccc tcg ctt ttc      2400
Glu Arg Arg Arg Leu Asp Glu Gln Gln Arg Ser Ala Pro Ser Leu Phe
785                 790                 795                 800

cgg cgc gcg tca gac gga cta aaa cgt cgc ttt agg gga tat aaa ccg      2448
Arg Arg Ala Ser Asp Gly Leu Lys Arg Arg Phe Arg Gly Tyr Lys Pro
                805                 810                 815

ctg gaa aac gaa gag gct caa gag tat gaa atg agc aaa taa              2490
Leu Glu Asn Glu Glu Ala Gln Glu Tyr Glu Met Ser Lys
            820                 825                 830


<210> 13
<211> 829
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 13
Met Met Ile Thr Asn Arg Thr Arg Arg Leu Leu Arg Ala Trp Val Val
 1               5                  10                  15
Ile Ile Ala Ile Gly Thr Ala Val Gly Glu Asn Val Thr Thr Pro Lys
            20                  25                  30
Gly Ala Thr Thr Thr Ala Lys Pro Thr Pro Gly Pro Ser Thr Pro Thr
        35                  40                  45
Pro Pro Glu Asn Pro Pro Arg Ala Glu Ala Phe Lys Phe Arg Val Cys
    50                  55                  60
Ser Ala Ser Ala Thr Gly Glu Leu Phe Arg Phe Asn Leu Glu Lys Thr
 65                 70                  75                  80
Cys Pro Gly Thr Glu Asp Lys Thr His Gln Glu Gly Ile Leu Met Val
                85                  90                  95
Phe Lys Lys Asn Ile Val Pro His Ile Phe Lys Val Arg Arg Tyr Arg
            100                 105                 110
Lys Val Ala Thr Ser Val Thr Val Tyr Arg Gly Trp Thr Glu Thr Ala
        115                 120                 125
Val Thr Gly Lys Gln Glu Val Ile Arg Pro Val Pro Gln Tyr Glu Ile
    130                 135                 140
Asn His Met Asp Thr Thr Tyr Gln Cys Phe Ser Ser Met Arg Val Asn
145                 150                 155                 160
Val Asn Gly Ile Val Asn Thr Tyr Thr Asp Arg Asp Phe Thr Asn Gln
                165                 170                 175
Thr Val Phe Leu Gln Pro Val Glu Gly Leu Thr Asp Asn Ile Gln Arg
            180                 185                 190
Tyr Phe Ser Gln Pro Val Leu Tyr Thr Thr Pro Gly Trp Phe Pro Gly
        195                 200                 205
Ile Tyr Arg Val Arg Thr Thr Val Asn Cys Glu Ile Val Asp Met Ile
    210                 215                 220
Ala Arg Ser Ala Glu Pro Tyr Ser Tyr Phe Val Thr Ala Leu Gly Asp
225                 230                 235                 240
Thr Val Glu Val Ser Pro Phe Cys His Asn Asp Ser Thr Cys Ser Val
                245                 250                 255
Ala Glu Lys Thr Glu Asn Gly Leu Gly Ala Arg Val Leu Thr Asn Tyr
            260                 265                 270
Thr Ile Val Asp Phe Ala Thr Arg Gln Pro Thr Thr Glu Thr Arg Val
        275                 280                 285
```

Phe Ala Asp Ser Gly Glu Tyr Thr Val Ser Trp Lys Ala Glu Asp Pro
290 295 300
Lys Ser Ala Val Cys Ala Leu Thr Leu Trp Lys Thr Phe Pro Arg Ala
305 310 315 320
Ile Gln Thr Thr His Glu Ala Ser Tyr His Phe Val Ala Asn Asp Val
325 330 335
Thr Ala Thr Phe Thr Ser Pro Leu Ser Gln Val Thr Asn Phe Thr Gly
340 345 350
Thr Tyr Pro Cys Leu Asn Asp Val Ile Gln Lys Thr Leu Asn Ala Thr
355 360 365
Ile Lys Lys Leu Ser Asp Thr His Ala Thr Asn Gly Ser Glu Gln Tyr
370 375 380
Tyr Glu Thr Glu Gly Gly Leu Phe Leu Leu Trp Gln Pro Leu Thr Pro
385 390 395 400
Leu Ser Leu Ala Asp Glu Met Arg Glu Leu Asn Gly Thr Thr Pro Ala
405 410 415
Pro Pro Thr Thr Thr Ser Thr Ala Asn Arg Val Arg Arg Ser Val Gly
420 425 430
Thr Asn Glu Gln Ala Thr Asp Asp Leu Ala Ala Pro Gln Leu Gln Phe
435 440 445
Ala Tyr Asp Lys Leu Arg Ala Ser Ile Asn Lys Val Leu Glu Glu Leu
450 455 460
Ser Arg Ala Trp Cys Arg Glu Gln Val Arg Asp Thr Tyr Met Trp Tyr
465 470 475 480
Glu Leu Ser Lys Ile Asn Pro Thr Ser Val Met Thr Ala Ile Tyr Gly
485 490 495
Arg Pro Val Ser Ala Lys Phe Val Gly Asp Ala Ile Ser Val Thr Asp
500 505 510
Cys Val Ala Val Asp Gln Ala Ser Val Ser Ile His Lys Ser Leu Arg
515 520 525
Thr Ser Thr Pro Gly Ile Cys Tyr Ser Arg Pro Pro Val Thr Phe Arg
530 535 540
Phe Leu Asn Ser Thr Thr Leu Phe Lys Gly Gln Leu Gly Pro Arg Asn
545 550 555 560
Glu Ile Ile Leu Thr Asp Asn Gln Val Glu Ala Cys Lys Glu Thr Cys
565 570 575
Glu His Tyr Phe Ile Ala Ser Asn Val Thr Tyr Tyr Tyr Lys Asp Tyr
580 585 590
Val Phe Val Lys Lys Ile Asn Thr Ser Glu Ile Ser Thr Leu Gly Thr
595 600 605
Phe Ile Ala Leu Asn Leu Ser Phe Ile Glu Asn Ile Asp Phe Arg Val
610 615 620
Ile Glu Leu Tyr Ser Arg Ala Glu Lys Lys Leu Ser Gly Ser Val Phe
625 630 635 640
Asp Ile Glu Thr Met Phe Arg Glu Tyr Asn Tyr Tyr Thr Gln Arg Leu
645 650 655
Ala Gly Leu Arg Glu Asp Leu Asp Asn Thr Ile Asp Leu Asn Arg Asp
660 665 670
Arg Leu Ala Arg Asp Leu Ser Glu Ile Val Ala Asp Leu Gly Asp Val
675 680 685
Gly Arg Thr Val Val Asn Val Ala Ser Ser Val Ile Thr Leu Phe Gly
690 695 700
Ser Ile Val Ser Gly Phe Ile Asn Phe Ile Lys Ser Pro Phe Gly Gly
705 710 715 720
Met Leu Met Ile Leu Val Ile Val Ala Val Val Leu Ile Val Phe Ala
725 730 735
Leu Asn Arg Arg Thr Asn Ala Ile Ala Gln Ala Pro Ile Arg Met Ile
740 745 750
Tyr Pro Asp Ile Asp Lys Met Gln Pro Ser Gly Gly Lys Val Asp Gln
755 760 765
Glu Gln Ile Lys Asn Ile Leu Ala Gly Met His Gln Leu Gln Gln Glu

770 775 780
Glu Arg Arg Arg Leu Asp Glu Gln Gln Arg Ser Ala Pro Ser Leu Phe
785 790 795 800
Arg Arg Ala Ser Asp Gly Leu Lys Arg Arg Phe Arg Gly Tyr Lys Pro
805 810 815
Leu Glu Asn Glu Glu Ala Gln Glu Tyr Glu Met Ser Lys
820 825

<210> 14
<211> 3045
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(3045)

<400> 14
atg gat ttc ttt aac ccg tac ctg ggc cct cgc gga cca cgc ccc cac 48
Met Asp Phe Phe Asn Pro Tyr Leu Gly Pro Arg Gly Pro Arg Pro His
1 5 10 15

tca cac aga ggc acc gat gct ccc gcc cct gcc ggc gcc gga gcc gtt 96
Ser His Arg Gly Thr Asp Ala Pro Ala Pro Ala Gly Ala Gly Ala Val
20 25 30

cag ccg cca cca gac gtt tgc agg ctc atc ccc gcc tgc ctc cga acg 144
Gln Pro Pro Pro Asp Val Cys Arg Leu Ile Pro Ala Cys Leu Arg Thr
35 40 45

cca ggg gca ggc ggg atg atc ccg gtc acg atc ccg ttc ccg cca acg 192
Pro Gly Ala Gly Gly Met Ile Pro Val Thr Ile Pro Phe Pro Pro Thr
50 55 60

tac ttc gag aac ggt gct cgc gga gac gtg ctg ctc gcc aac gaa cgg 240
Tyr Phe Glu Asn Gly Ala Arg Gly Asp Val Leu Leu Ala Asn Glu Arg
65 70 75 80

tcc atg tgg acg gcg cgc gac cgc aag ccc gtc gcc ccg gac ccc caa 288
Ser Met Trp Thr Ala Arg Asp Arg Lys Pro Val Ala Pro Asp Pro Gln
85 90 95

gac caa tcc atc acg ttt cac gcg tac gac gtc gtt gaa aca acg tac 336
Asp Gln Ser Ile Thr Phe His Ala Tyr Asp Val Val Glu Thr Thr Tyr
100 105 110

gcg gcg gac agg tgt gcc gag gta cct agc cgc ttc caa acg gac att 384
Ala Ala Asp Arg Cys Ala Glu Val Pro Ser Arg Phe Gln Thr Asp Ile
115 120 125

atc cca agc gga acc gtg ctc aag ctc ctg ggg cga acc gag gac ggc 432
Ile Pro Ser Gly Thr Val Leu Lys Leu Leu Gly Arg Thr Glu Asp Gly
130 135 140

acc agc gtg tgc gtg aac gtg ttc cgt caa cag gta tat ttc tac gcg 480
Thr Ser Val Cys Val Asn Val Phe Arg Gln Gln Val Tyr Phe Tyr Ala
145 150 155 160

aag gtt cca gcc ggc att aac gtc acc cac atc ctc cag cag gcc ctc 528

```
Lys Val Pro Ala Gly Ile Asn Val Thr His Ile Leu Gln Gln Ala Leu
            165                 170                 175

aag aac aca gcc ggc cgt gcc gcg tgc ggc ttc tcg acc aga aga gtg     576
Lys Asn Thr Ala Gly Arg Ala Ala Cys Gly Phe Ser Thr Arg Arg Val
        180                 185                 190

aac aaa aga att ctc aaa acg tac gac gtc gcg gag cat ccc gtc acg     624
Asn Lys Arg Ile Leu Lys Thr Tyr Asp Val Ala Glu His Pro Val Thr
    195                 200                 205

gaa atc acg cta tcg tcc ggt tcc atg ctc tcg acc ctc agc gac cgc     672
Glu Ile Thr Leu Ser Ser Gly Ser Met Leu Ser Thr Leu Ser Asp Arg
210                 215                 220

ctc gtc gcg tgc ggg tgc gag gtg ttc gag tca aac gtg gac gcc gtt     720
Leu Val Ala Cys Gly Cys Glu Val Phe Glu Ser Asn Val Asp Ala Val
225                 230                 235                 240

cgc cgg ttc gtt ctg gat cac ggg ttt acc acg ttc ggg tgg tac tcg     768
Arg Arg Phe Val Leu Asp His Gly Phe Thr Thr Phe Gly Trp Tyr Ser
            245                 250                 255

tgc gcg cgc gcc acg ccc cgc ctg gcg gcc aga gat gcc agg acg gcc     816
Cys Ala Arg Ala Thr Pro Arg Leu Ala Ala Arg Asp Ala Arg Thr Ala
        260                 265                 270

ctg gag ttt gac tgc agc tgg gag gac ctc agc gtt caa gcg gac cgc     864
Leu Glu Phe Asp Cys Ser Trp Glu Asp Leu Ser Val Gln Ala Asp Arg
    275                 280                 285

agc gac tgg ccc ccg tac cgc atc gtg gcc ttt gat atc gag tgc act     912
Ser Asp Trp Pro Pro Tyr Arg Ile Val Ala Phe Asp Ile Glu Cys Thr
290                 295                 300

gga gag gcg gga ttt ccg tgc gcc acg cgc gac ggc gac gcg gtg atc     960
Gly Glu Ala Gly Phe Pro Cys Ala Thr Arg Asp Gly Asp Ala Val Ile
305                 310                 315                 320

cag atc tcc tgc gtc ttc tac acg acc agg gaa ggc gcg ccc aat ccg    1008
Gln Ile Ser Cys Val Phe Tyr Thr Thr Arg Glu Gly Ala Pro Asn Pro
            325                 330                 335

cca aac ata ctg ttc agc gtc ggg acg tgc gac ccc atc ccg gac acc    1056
Pro Asn Ile Leu Phe Ser Val Gly Thr Cys Asp Pro Ile Pro Asp Thr
        340                 345                 350

gac gtt ttg gag ttt ccg tcg gaa tat gac atg ctg gtg tcg ttc ttc    1104
Asp Val Leu Glu Phe Pro Ser Glu Tyr Asp Met Leu Val Ser Phe Phe
    355                 360                 365

gcc atg atc cgc gac ttc gag gtg gac ttt tta acc ggc tat aac atc    1152
Ala Met Ile Arg Asp Phe Glu Val Asp Phe Leu Thr Gly Tyr Asn Ile
370                 375                 380

tca aac ttc gat ctc ccg tac cta atc acg cga gcg tcc cag gtg tac    1200
Ser Asn Phe Asp Leu Pro Tyr Leu Ile Thr Arg Ala Ser Gln Val Tyr
385                 390                 395                 400

aac ctt cga tta aac gaa tac aca aaa ata aaa acc ggc tcc atc ttt    1248
Asn Leu Arg Leu Asn Glu Tyr Thr Lys Ile Lys Thr Gly Ser Ile Phe
```

405 410 415

gaa gtt cac gag ccc cgt ggc ggg gga ggg ggg ttc atg agg tcg gtc 1296
Glu Val His Glu Pro Arg Gly Gly Gly Gly Gly Phe Met Arg Ser Val
420 425 430

tca aaa att aaa ata gcg ggc atc gtc ccc ata gac atg tac cag gtg 1344
Ser Lys Ile Lys Ile Ala Gly Ile Val Pro Ile Asp Met Tyr Gln Val
435 440 445

tgt cgc gaa aag ctc agc ctc tcc gac tac aaa ctg gac acg gtg gcc 1392
Cys Arg Glu Lys Leu Ser Leu Ser Asp Tyr Lys Leu Asp Thr Val Ala
450 455 460

agg cag tgt ctg ggt ggg aaa aaa gag gac gta tcg tac aag gac att 1440
Arg Gln Cys Leu Gly Gly Lys Lys Glu Asp Val Ser Tyr Lys Asp Ile
465 470 475 480

ccc cct ctg ttt cgc tca ggt ccg ggc ggc agg gct aag gtg ggc agc 1488
Pro Pro Leu Phe Arg Ser Gly Pro Gly Gly Arg Ala Lys Val Gly Ser
485 490 495

tat tgc gtg atg gac tcg gtc ctg gtg atg gac ctc tta aaa atg ttt 1536
Tyr Cys Val Met Asp Ser Val Leu Val Met Asp Leu Leu Lys Met Phe
500 505 510

atg ata cac gtg gag att tcg gag ata gcc aag ctg gcc aag att cag 1584
Met Ile His Val Glu Ile Ser Glu Ile Ala Lys Leu Ala Lys Ile Gln
515 520 525

gcc agg cgc gtc ctg acg gac ggc caa cag ctc cgc gtg ttc tcc tgc 1632
Ala Arg Arg Val Leu Thr Asp Gly Gln Gln Leu Arg Val Phe Ser Cys
530 535 540

ctg ctg gag gcc gcg gcc agg gag aac ttt atc ctc ccg gtt cca acg 1680
Leu Leu Glu Ala Ala Ala Arg Glu Asn Phe Ile Leu Pro Val Pro Thr
545 550 555 560

ccc gag gga cag ggg ggc tat cag ggc gcg acg gtg atc aac ccc att 1728
Pro Glu Gly Gln Gly Gly Tyr Gln Gly Ala Thr Val Ile Asn Pro Ile
565 570 575

ccg ggg ttt tac gac gag ccg gtc ctg gtg gtc gat ttt gcc agc ctg 1776
Pro Gly Phe Tyr Asp Glu Pro Val Leu Val Val Asp Phe Ala Ser Leu
580 585 590

tac ccg agc atc atc cag gcg cac aac ctg tgc tac tcc acc atg ata 1824
Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Tyr Ser Thr Met Ile
595 600 605

cac gga cga gac ctg cac ctg cac ccc aac ctg acg ccg gac gac tac 1872
His Gly Arg Asp Leu His Leu His Pro Asn Leu Thr Pro Asp Asp Tyr
610 615 620

gag acg ttc gtg ctg agc ggc gga ccg gta cat ttt gta aaa aaa cac 1920
Glu Thr Phe Val Leu Ser Gly Gly Pro Val His Phe Val Lys Lys His
625 630 635 640

aag cgg gag tct ctg ctg gga aga ctg cta acc gtg tgg tta gaa aag 1968
Lys Arg Glu Ser Leu Leu Gly Arg Leu Leu Thr Val Trp Leu Glu Lys
645 650 655

```
cga agg gcg atc cgg cgc acc ctg gcg gcg tgc gat gac ccg tcg cta    2016
Arg Arg Ala Ile Arg Arg Thr Leu Ala Ala Cys Asp Asp Pro Ser Leu
            660                 665                 670

aaa acc atc tta gat aaa caa cag ctg gcc atc aag gtg aca tgt aac    2064
Lys Thr Ile Leu Asp Lys Gln Gln Leu Ala Ile Lys Val Thr Cys Asn
        675                 680                 685

gcg gtt tac ggg ttc acc ggg gtg gcc agc ggc ctc ctc cca tgc att    2112
Ala Val Tyr Gly Phe Thr Gly Val Ala Ser Gly Leu Leu Pro Cys Ile
    690                 695                 700

aac ata gcg gaa acc gtg acg ctc cgg ggg cgc acg atg ctg gag atg    2160
Asn Ile Ala Glu Thr Val Thr Leu Arg Gly Arg Thr Met Leu Glu Met
705                 710                 715                 720

tca aag tct tac gtg gag gcc ctg acg acg gaa gac ctg cga acg cgt    2208
Ser Lys Ser Tyr Val Glu Ala Leu Thr Thr Glu Asp Leu Arg Thr Arg
            725                 730                 735

ctc ggt cgc gag gtg acc gcc cgt cac ggc gcg cgg ttt cgc gtc gtc    2256
Leu Gly Arg Glu Val Thr Ala Arg His Gly Ala Arg Phe Arg Val Val
            740                 745                 750

tac ggt gac acc gac tcc ctc ttt atc gcg tgc gac ggt tat tcc gcg    2304
Tyr Gly Asp Thr Asp Ser Leu Phe Ile Ala Cys Asp Gly Tyr Ser Ala
        755                 760                 765

gaa gcc gtt tcc gct ttc tgt gac gat ctg gcc gcc agg atc act gcg    2352
Glu Ala Val Ser Ala Phe Cys Asp Asp Leu Ala Ala Arg Ile Thr Ala
    770                 775                 780

gac ctg ttc ccc cca ccc att aag cta gag gcg gaa aag acg ttc aag    2400
Asp Leu Phe Pro Pro Pro Ile Lys Leu Glu Ala Glu Lys Thr Phe Lys
785                 790                 795                 800

tgt ctg ctg ctg ctg acg aaa aag cgc tac atc ggg gtc cta ttg aac    2448
Cys Leu Leu Leu Leu Thr Lys Lys Arg Tyr Ile Gly Val Leu Leu Asn
            805                 810                 815

gac aaa atg gtc atg aaa ggg gtc gac ctc att cgc aaa acg gcc tgc    2496
Asp Lys Met Val Met Lys Gly Val Asp Leu Ile Arg Lys Thr Ala Cys
            820                 825                 830

aag ttt gtc cag gag cga tgc cgc gcc atc ctg gac ctg gtg ctc cac    2544
Lys Phe Val Gln Glu Arg Cys Arg Ala Ile Leu Asp Leu Val Leu His
        835                 840                 845

gat ccg gag gtc aag gct gcg gcg cgg ctg ttg tgc aag cgg ccg ccg    2592
Asp Pro Glu Val Lys Ala Ala Ala Arg Leu Leu Cys Lys Arg Pro Pro
    850                 855                 860

cac gcg gta tac gag gag ggg ctg ccg gct ggc ttt ata aaa atc gta    2640
His Ala Val Tyr Glu Glu Gly Leu Pro Ala Gly Phe Ile Lys Ile Val
865                 870                 875                 880

gag gtc ctc aac gcg agc tat ctg gac ctc cga aac agc gtc gtg ccc    2688
Glu Val Leu Asn Ala Ser Tyr Leu Asp Leu Arg Asn Ser Val Val Pro
            885                 890                 895
```

atc gag cag tta acg ttc tcc acc gag ctc agc cgc ccc gtc tgc gat 2736
Ile Glu Gln Leu Thr Phe Ser Thr Glu Leu Ser Arg Pro Val Cys Asp
900 905 910

tac aag acc acc aac ctg ccc cac ctg gcg gtg tac caa aag ctg gcg 2784
Tyr Lys Thr Thr Asn Leu Pro His Leu Ala Val Tyr Gln Lys Leu Ala
915 920 925

agc agg tgc gag gag ctg ccc cag gtg cac gat aga atc ccc tac gtg 2832
Ser Arg Cys Glu Glu Leu Pro Gln Val His Asp Arg Ile Pro Tyr Val
930 935 940

ttc gtt gac gcg ccc ggg tcc cta aag tcg gac ctg gcc gaa cac ccg 2880
Phe Val Asp Ala Pro Gly Ser Leu Lys Ser Asp Leu Ala Glu His Pro
945 950 955 960

gat tac gtc aga cag cac cag att ccc gtc gcg gtc gac cta tat ttc 2928
Asp Tyr Val Arg Gln His Gln Ile Pro Val Ala Val Asp Leu Tyr Phe
965 970 975

gac aaa ctg gtg cac ggc gcg gcc aac atc ctc cag tgt ctg ttc ggc 2976
Asp Lys Leu Val His Gly Ala Ala Asn Ile Leu Gln Cys Leu Phe Gly
980 985 990

aac aac gcg gac acc acg gtg gcc atc ctc tac aat ttt ctc aac gtc 3024
Asn Asn Ala Asp Thr Thr Val Ala Ile Leu Tyr Asn Phe Leu Asn Val
995 1000 1005

ccg tat aag ctg ttc tcg tga 3045
Pro Tyr Lys Leu Phe Ser
1010 1015

<210> 15
<211> 1014
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 15
Met Asp Phe Phe Asn Pro Tyr Leu Gly Pro Arg Gly Pro Arg Pro His
1 5 10 15
Ser His Arg Gly Thr Asp Ala Pro Ala Pro Ala Gly Ala Gly Ala Val
20 25 30
Gln Pro Pro Pro Asp Val Cys Arg Leu Ile Pro Ala Cys Leu Arg Thr
35 40 45
Pro Gly Ala Gly Gly Met Ile Pro Val Thr Ile Pro Phe Pro Pro Thr
50 55 60
Tyr Phe Glu Asn Gly Ala Arg Gly Asp Val Leu Leu Ala Asn Glu Arg
65 70 75 80
Ser Met Trp Thr Ala Arg Asp Arg Lys Pro Val Ala Pro Asp Pro Gln
85 90 95
Asp Gln Ser Ile Thr Phe His Ala Tyr Asp Val Val Glu Thr Thr Tyr
100 105 110
Ala Ala Asp Arg Cys Ala Glu Val Pro Ser Arg Phe Gln Thr Asp Ile
115 120 125
Ile Pro Ser Gly Thr Val Leu Lys Leu Leu Gly Arg Thr Glu Asp Gly
130 135 140
Thr Ser Val Cys Val Asn Val Phe Arg Gln Gln Val Tyr Phe Tyr Ala
145 150 155 160
Lys Val Pro Ala Gly Ile Asn Val Thr His Ile Leu Gln Gln Ala Leu
165 170 175

Lys Asn Thr Ala Gly Arg Ala Ala Cys Gly Phe Ser Thr Arg Arg Val
180 185 190
Asn Lys Arg Ile Leu Lys Thr Tyr Asp Val Ala Glu His Pro Val Thr
195 200 205
Glu Ile Thr Leu Ser Ser Gly Ser Met Leu Ser Thr Leu Ser Asp Arg
210 215 220
Leu Val Ala Cys Gly Cys Glu Val Phe Glu Ser Asn Val Asp Ala Val
225 230 235 240
Arg Arg Phe Val Leu Asp His Gly Phe Thr Thr Phe Gly Trp Tyr Ser
245 250 255
Cys Ala Arg Ala Thr Pro Arg Leu Ala Ala Arg Asp Ala Arg Thr Ala
260 265 270
Leu Glu Phe Asp Cys Ser Trp Glu Asp Leu Ser Val Gln Ala Asp Arg
275 280 285
Ser Asp Trp Pro Pro Tyr Arg Ile Val Ala Phe Asp Ile Glu Cys Thr
290 295 300
Gly Glu Ala Gly Phe Pro Cys Ala Thr Arg Asp Gly Asp Ala Val Ile
305 310 315 320
Gln Ile Ser Cys Val Phe Tyr Thr Thr Arg Glu Gly Ala Pro Asn Pro
325 330 335
Pro Asn Ile Leu Phe Ser Val Gly Thr Cys Asp Pro Ile Pro Asp Thr
340 345 350
Asp Val Leu Glu Phe Pro Ser Glu Tyr Asp Met Leu Val Ser Phe Phe
355 360 365
Ala Met Ile Arg Asp Phe Glu Val Asp Phe Leu Thr Gly Tyr Asn Ile
370 375 380
Ser Asn Phe Asp Leu Pro Tyr Leu Ile Thr Arg Ala Ser Gln Val Tyr
385 390 395 400
Asn Leu Arg Leu Asn Glu Tyr Thr Lys Ile Lys Thr Gly Ser Ile Phe
405 410 415
Glu Val His Glu Pro Arg Gly Gly Gly Gly Gly Phe Met Arg Ser Val
420 425 430
Ser Lys Ile Lys Ile Ala Gly Ile Val Pro Ile Asp Met Tyr Gln Val
435 440 445
Cys Arg Glu Lys Leu Ser Leu Ser Asp Tyr Lys Leu Asp Thr Val Ala
450 455 460
Arg Gln Cys Leu Gly Gly Lys Lys Glu Asp Val Ser Tyr Lys Asp Ile
465 470 475 480
Pro Pro Leu Phe Arg Ser Gly Pro Gly Gly Arg Ala Lys Val Gly Ser
485 490 495
Tyr Cys Val Met Asp Ser Val Leu Val Met Asp Leu Leu Lys Met Phe
500 505 510
Met Ile His Val Glu Ile Ser Glu Ile Ala Lys Leu Ala Lys Ile Gln
515 520 525
Ala Arg Arg Val Leu Thr Asp Gly Gln Gln Leu Arg Val Phe Ser Cys
530 535 540
Leu Leu Glu Ala Ala Ala Arg Glu Asn Phe Ile Leu Pro Val Pro Thr
545 550 555 560
Pro Glu Gly Gln Gly Gly Tyr Gln Gly Ala Thr Val Ile Asn Pro Ile
565 570 575
Pro Gly Phe Tyr Asp Glu Pro Val Leu Val Val Asp Phe Ala Ser Leu
580 585 590
Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Tyr Ser Thr Met Ile
595 600 605
His Gly Arg Asp Leu His Leu His Pro Asn Leu Thr Pro Asp Asp Tyr
610 615 620
Glu Thr Phe Val Leu Ser Gly Gly Pro Val His Phe Val Lys Lys His
625 630 635 640
Lys Arg Glu Ser Leu Leu Gly Arg Leu Leu Thr Val Trp Leu Glu Lys
645 650 655
Arg Arg Ala Ile Arg Arg Thr Leu Ala Ala Cys Asp Asp Pro Ser Leu

660 665 670
Lys Thr Ile Leu Asp Lys Gln Gln Leu Ala Ile Lys Val Thr Cys Asn
675 680 685
Ala Val Tyr Gly Phe Thr Gly Val Ala Ser Gly Leu Leu Pro Cys Ile
690 695 700
Asn Ile Ala Glu Thr Val Thr Leu Arg Gly Arg Thr Met Leu Glu Met
705 710 715 720
Ser Lys Ser Tyr Val Glu Ala Leu Thr Thr Glu Asp Leu Arg Thr Arg
725 730 735
Leu Gly Arg Glu Val Thr Ala Arg His Gly Ala Arg Phe Arg Val Val
740 745 750
Tyr Gly Asp Thr Asp Ser Leu Phe Ile Ala Cys Asp Gly Tyr Ser Ala
755 760 765
Glu Ala Val Ser Ala Phe Cys Asp Asp Leu Ala Ala Arg Ile Thr Ala
770 775 780
Asp Leu Phe Pro Pro Pro Ile Lys Leu Glu Ala Glu Lys Thr Phe Lys
785 790 795 800
Cys Leu Leu Leu Leu Thr Lys Lys Arg Tyr Ile Gly Val Leu Leu Asn
805 810 815
Asp Lys Met Val Met Lys Gly Val Asp Leu Ile Arg Lys Thr Ala Cys
820 825 830
Lys Phe Val Gln Glu Arg Cys Arg Ala Ile Leu Asp Leu Val Leu His
835 840 845
Asp Pro Glu Val Lys Ala Ala Ala Arg Leu Leu Cys Lys Arg Pro Pro
850 855 860
His Ala Val Tyr Glu Glu Gly Leu Pro Ala Gly Phe Ile Lys Ile Val
865 870 875 880
Glu Val Leu Asn Ala Ser Tyr Leu Asp Leu Arg Asn Ser Val Val Pro
885 890 895
Ile Glu Gln Leu Thr Phe Ser Thr Glu Leu Ser Arg Pro Val Cys Asp
900 905 910
Tyr Lys Thr Thr Asn Leu Pro His Leu Ala Val Tyr Gln Lys Leu Ala
915 920 925
Ser Arg Cys Glu Glu Leu Pro Gln Val His Asp Arg Ile Pro Tyr Val
930 935 940
Phe Val Asp Ala Pro Gly Ser Leu Lys Ser Asp Leu Ala Glu His Pro
945 950 955 960
Asp Tyr Val Arg Gln His Gln Ile Pro Val Ala Val Asp Leu Tyr Phe
965 970 975
Asp Lys Leu Val His Gly Ala Ala Asn Ile Leu Gln Cys Leu Phe Gly
980 985 990
Asn Asn Ala Asp Thr Thr Val Ala Ile Leu Tyr Asn Phe Leu Asn Val
995 1000 1005
Pro Tyr Lys Leu Phe Ser
1010

<210> 16
<211> 1251
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1251)

<400> 16
atg ctg gtt aac gaa ctg tcg gtg gtc ctc ggc gac tgg gag gtg act 48
Met Leu Val Asn Glu Leu Ser Val Val Leu Gly Asp Trp Glu Val Thr
1 5 10 15

```
ttt cac cgg ggt aga ttc agc ttc gtc aac ctc acc cgc ctg caa acg     96
Phe His Arg Gly Arg Phe Ser Phe Val Asn Leu Thr Arg Leu Gln Thr
            20                  25                  30

ttc aag ggc cac ggg ggc tac gcc agg gtc cga ctc ccc ttc tcg ctc     144
Phe Lys Gly His Gly Gly Tyr Ala Arg Val Arg Leu Pro Phe Ser Leu
        35                  40                  45

gac cag tta ctc cac caa cat ttc gcg ttc gga ctc gtg acg cgt ctc     192
Asp Gln Leu Leu His Gln His Phe Ala Phe Gly Leu Val Thr Arg Leu
    50                  55                  60

aag gaa ctg ccc ccc ttc tcc gac tgc gtg gcc ctt atc gcc ccg ttg     240
Lys Glu Leu Pro Pro Phe Ser Asp Cys Val Ala Leu Ile Ala Pro Leu
65                  70                  75                  80

gat tcc ggc ggc gac gcg gac gcg gcg cgc gtg gcc ccc ggg ttc gtg     288
Asp Ser Gly Gly Asp Ala Asp Ala Ala Arg Val Ala Pro Gly Phe Val
                85                  90                  95

ctg gac tcc tct cgc ccg ctg acc gtg tgg gta aac gcg agc ggg cgg     336
Leu Asp Ser Ser Arg Pro Leu Thr Val Trp Val Asn Ala Ser Gly Arg
            100                 105                 110

cac acg atc cgg ttc tgc ctc ctc ttt cta aag ccg atc gac ctg gag     384
His Thr Ile Arg Phe Cys Leu Leu Phe Leu Lys Pro Ile Asp Leu Glu
        115                 120                 125

cgc gcg gtc acg tac gtc ttc ggc gag aac ggc ggc gcg cgc tcg gag     432
Arg Ala Val Thr Tyr Val Phe Gly Glu Asn Gly Gly Ala Arg Ser Glu
    130                 135                 140

ggc acc cca aag ccc acc tgc gcg acc gaa agc ctg ccc ggt gga ccc     480
Gly Thr Pro Lys Pro Thr Cys Ala Thr Glu Ser Leu Pro Gly Gly Pro
145                 150                 155                 160

ctg cgc gtc tcc ggc gag gcg tct cag acg tcg ccc cat tct ttc gtt     528
Leu Arg Val Ser Gly Glu Ala Ser Gln Thr Ser Pro His Ser Phe Val
                165                 170                 175

gcg tat ttt ccc acg gcc aac tcg gtg gcc tgc cta agc ctg ttg cgg     576
Ala Tyr Phe Pro Thr Ala Asn Ser Val Ala Cys Leu Ser Leu Leu Arg
            180                 185                 190

tta cag gtg agg ccg ttt tcg gat gac gcg gcg cac agg gac gcg cgg     624
Leu Gln Val Arg Pro Phe Ser Asp Asp Ala Ala His Arg Asp Ala Arg
        195                 200                 205

atc tcc ccg aaa tac gtc acg ttt agt aac tcc ggg ggt aac gtc tgc     672
Ile Ser Pro Lys Tyr Val Thr Phe Ser Asn Ser Gly Gly Asn Val Cys
    210                 215                 220

aag gcg tcc gtt cac acg ctg tcc ccg tcg cgg tgt aaa acg gcg caa     720
Lys Ala Ser Val His Thr Leu Ser Pro Ser Arg Cys Lys Thr Ala Gln
225                 230                 235                 240

atg gaa atc atc tac gct ccc ggg gac ccc aac gcc gag ata gtc ctg     768
Met Glu Ile Ile Tyr Ala Pro Gly Asp Pro Asn Ala Glu Ile Val Leu
                245                 250                 255
```

ggc cag tcc gga ccc gtc ctg ccc acc cac acc ggc ggc cgc gta ttg 816
Gly Gln Ser Gly Pro Val Leu Pro Thr His Thr Gly Gly Arg Val Leu
260 265 270

ggg gtc tac gcc gac gcc gaa aaa acc atc caa cct gga agc tcc gcg 864
Gly Val Tyr Ala Asp Ala Glu Lys Thr Ile Gln Pro Gly Ser Ser Ala
275 280 285

gaa gtc cgg gtt cag tta atc ttc caa cag gga gcg gcc gct cgg ggc 912
Glu Val Arg Val Gln Leu Ile Phe Gln Gln Gly Ala Ala Ala Arg Gly
290 295 300

gat ctg gcg ttt ctg gtc acg ggc gtg gca ccg gag ccc cta ttc gtc 960
Asp Leu Ala Phe Leu Val Thr Gly Val Ala Pro Glu Pro Leu Phe Val
305 310 315 320

gtc acc ccg gca ctc ttg ctt tcc ggt tgc aca acc cac ctg cgc cta 1008
Val Thr Pro Ala Leu Leu Leu Ser Gly Cys Thr Thr His Leu Arg Leu
325 330 335

ttc aac ccc aac ggt acc ccc acg act ata aaa aga gac acc ctt gtg 1056
Phe Asn Pro Asn Gly Thr Pro Thr Thr Ile Lys Arg Asp Thr Leu Val
340 345 350

gcc gcc gcc gcg ccc tgc ccc gtg gtg cga tta agc tcc gcc gac gac 1104
Ala Ala Ala Ala Pro Cys Pro Val Val Arg Leu Ser Ser Ala Asp Asp
355 360 365

gcg ccg cga gac ctc gtc gcg tca cca gac acc ggg gcg ctc tcc att 1152
Ala Pro Arg Asp Leu Val Ala Ser Pro Asp Thr Gly Ala Leu Ser Ile
370 375 380

aac gcg ttc aca atc ccg gtc ggt ttc cca ggg gtg gtc tcg gcg gag 1200
Asn Ala Phe Thr Ile Pro Val Gly Phe Pro Gly Val Val Ser Ala Glu
385 390 395 400

tgt cac gtg tcc cta cgc gac aac ggg gtc cac gaa cgc atg aac cat 1248
Cys His Val Ser Leu Arg Asp Asn Gly Val His Glu Arg Met Asn His
405 410 415

tga 1251

<210> 17
<211> 416
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 17
Met Leu Val Asn Glu Leu Ser Val Val Leu Gly Asp Trp Glu Val Thr
1 5 10 15
Phe His Arg Gly Arg Phe Ser Phe Val Asn Leu Thr Arg Leu Gln Thr
20 25 30
Phe Lys Gly His Gly Gly Tyr Ala Arg Val Arg Leu Pro Phe Ser Leu
35 40 45
Asp Gln Leu Leu His Gln His Phe Ala Phe Gly Leu Val Thr Arg Leu
50 55 60
Lys Glu Leu Pro Pro Phe Ser Asp Cys Val Ala Leu Ile Ala Pro Leu
65 70 75 80
Asp Ser Gly Gly Asp Ala Asp Ala Ala Arg Val Ala Pro Gly Phe Val

```
                    85                  90                  95
Leu Asp Ser Ser Arg Pro Leu Thr Val Trp Val Asn Ala Ser Gly Arg
                100                 105                 110
His Thr Ile Arg Phe Cys Leu Leu Phe Leu Lys Pro Ile Asp Leu Glu
            115                 120                 125
Arg Ala Val Thr Tyr Val Phe Gly Glu Asn Gly Gly Ala Arg Ser Glu
    130                 135                 140
Gly Thr Pro Lys Pro Thr Cys Ala Thr Glu Ser Leu Pro Gly Gly Pro
145                 150                 155                 160
Leu Arg Val Ser Gly Glu Ala Ser Gln Thr Ser Pro His Ser Phe Val
                165                 170                 175
Ala Tyr Phe Pro Thr Ala Asn Ser Val Ala Cys Leu Ser Leu Leu Arg
                180                 185                 190
Leu Gln Val Arg Pro Phe Ser Asp Asp Ala Ala His Arg Asp Ala Arg
            195                 200                 205
Ile Ser Pro Lys Tyr Val Thr Phe Ser Asn Ser Gly Gly Asn Val Cys
    210                 215                 220
Lys Ala Ser Val His Thr Leu Ser Pro Ser Arg Cys Lys Thr Ala Gln
225                 230                 235                 240
Met Glu Ile Ile Tyr Ala Pro Gly Asp Pro Asn Ala Glu Ile Val Leu
                245                 250                 255
Gly Gln Ser Gly Pro Val Leu Pro Thr His Thr Gly Gly Arg Val Leu
                260                 265                 270
Gly Val Tyr Ala Asp Ala Glu Lys Thr Ile Gln Pro Gly Ser Ser Ala
            275                 280                 285
Glu Val Arg Val Gln Leu Ile Phe Gln Gln Gly Ala Ala Ala Arg Gly
    290                 295                 300
Asp Leu Ala Phe Leu Val Thr Gly Val Ala Pro Glu Pro Leu Phe Val
305                 310                 315                 320
Val Thr Pro Ala Leu Leu Leu Ser Gly Cys Thr Thr His Leu Arg Leu
                325                 330                 335
Phe Asn Pro Asn Gly Thr Pro Thr Thr Ile Lys Arg Asp Thr Leu Val
                340                 345                 350
Ala Ala Ala Ala Pro Cys Pro Val Val Arg Leu Ser Ser Ala Asp Asp
            355                 360                 365
Ala Pro Arg Asp Leu Val Ala Ser Pro Asp Thr Gly Ala Leu Ser Ile
    370                 375                 380
Asn Ala Phe Thr Ile Pro Val Gly Phe Pro Gly Val Val Ser Ala Glu
385                 390                 395                 400
Cys His Val Ser Leu Arg Asp Asn Gly Val His Glu Arg Met Asn His
                405                 410                 415


<210> 18
<211> 1230
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1230)

<400> 18
atg gga aca ccg gtg cgt ttc ttt cgc ggc gag tgg cag acc tcg agt      48
Met Gly Thr Pro Val Arg Phe Phe Arg Gly Glu Trp Gln Thr Ser Ser
  1               5                  10                  15

cta gtg gac aac ggc acg cca cgg tac agc tcc ctg gtg tgg gcc gcc      96
Leu Val Asp Asn Gly Thr Pro Arg Tyr Ser Ser Leu Val Trp Ala Ala
             20                  25                  30
```

```
act att cac gac ggc tac ctg aca ctg gtg aac agg tca gag ctg tgc    144
Thr Ile His Asp Gly Tyr Leu Thr Leu Val Asn Arg Ser Glu Leu Cys
         35                  40                  45

gtc acg gag agg tct ccg tgt ctg ccg gca tgc ccc agc atc ggg aga    192
Val Thr Glu Arg Ser Pro Cys Leu Pro Ala Cys Pro Ser Ile Gly Arg
     50                  55                  60

ctg gtc ggg aag agg ttt ccc ggc ttc gcc ttt gcc agc gcc act ctg    240
Leu Val Gly Lys Arg Phe Pro Gly Phe Ala Phe Ala Ser Ala Thr Leu
 65                  70                  75                  80

ggc gat cgg gga aca cgc acc gtg ttc tac gcg ttc ggt cac cgc gac    288
Gly Asp Arg Gly Thr Arg Thr Val Phe Tyr Ala Phe Gly His Arg Asp
                 85                  90                  95

aac cca ctg gac ata gta ccc gcc gtg gtc gag cgc gcg gat cgc gag    336
Asn Pro Leu Asp Ile Val Pro Ala Val Val Glu Arg Ala Asp Arg Glu
            100                 105                 110

ctg gtg ctg cgg gtt cac gct ccg caa aca acg cgg gtg tcg cga tac    384
Leu Val Leu Arg Val His Ala Pro Gln Thr Thr Arg Val Ser Arg Tyr
        115                 120                 125

gga ctt aag gta ttc gtg gcg atc gtt acg gtg gtg cgc ccg ccc ggg    432
Gly Leu Lys Val Phe Val Ala Ile Val Thr Val Val Arg Pro Pro Gly
    130                 135                 140

gtg ttc cta cac ttt cca caa gac cgc gtt ccg atc gcg ctg aca gac    480
Val Phe Leu His Phe Pro Gln Asp Arg Val Pro Ile Ala Leu Thr Asp
145                 150                 155                 160

gcg tgc agc cag gag ggc tcc agg cta acc tct gaa gag ccg tgg ata    528
Ala Cys Ser Gln Glu Gly Ser Arg Leu Thr Ser Glu Glu Pro Trp Ile
                165                 170                 175

aaa att caa ggc ttt ccc gtc cta tct gac gag acc gcg cac cca ttt    576
Lys Ile Gln Gly Phe Pro Val Leu Ser Asp Glu Thr Ala His Pro Phe
            180                 185                 190

ctc tta acc cag aag acc aag ccc ttt acc gag cga aag ttt tgc cgc    624
Leu Leu Thr Gln Lys Thr Lys Pro Phe Thr Glu Arg Lys Phe Cys Arg
        195                 200                 205

ctg atc atg gac aac gac cag cgc agc gcc gtc aac acc gtc tac ctg    672
Leu Ile Met Asp Asn Asp Gln Arg Ser Ala Val Asn Thr Val Tyr Leu
    210                 215                 220

gga aag cag cac gtg agg gtg acc gtg acc cgc ccc ccg gaa aca atc    720
Gly Lys Gln His Val Arg Val Thr Val Thr Arg Pro Pro Glu Thr Ile
225                 230                 235                 240

gtc acc gac ggc ccc gtg acg gcg acc ctg tcc ctc acc ggt aat gcg    768
Val Thr Asp Gly Pro Val Thr Ala Thr Leu Ser Leu Thr Gly Asn Ala
                245                 250                 255

cca atc gcc ttt cgc cac aac cca tac ttt gaa ctc ccg tgg tcg tcc    816
Pro Ile Ala Phe Arg His Asn Pro Tyr Phe Glu Leu Pro Trp Ser Ser
            260                 265                 270
```

```
aca acg gcg ata ttc acg ccc gtg gtg tac gtg ggc ctg acc gtg tgc      864
Thr Thr Ala Ile Phe Thr Pro Val Val Tyr Val Gly Leu Thr Val Cys
        275                 280                 285

atc cca ccc aac tgt agc aaa ttc gta agg tac ggt aac acc tac gtc      912
Ile Pro Pro Asn Cys Ser Lys Phe Val Arg Tyr Gly Asn Thr Tyr Val
    290                 295                 300

tcg gca ttt aac cgc aag ctg acg gcg att att agc aat cac gcc cac      960
Ser Ala Phe Asn Arg Lys Leu Thr Ala Ile Ile Ser Asn His Ala His
305                 310                 315                 320

aac ggc ggg ttc cgg att cag gac tgc gag tgg cca ccg aac cgg gag     1008
Asn Gly Gly Phe Arg Ile Gln Asp Cys Glu Trp Pro Pro Asn Arg Glu
                325                 330                 335

ata gag att ttg gta acc aac gtg tcc cag gcc ccg gtg tac atc agc     1056
Ile Glu Ile Leu Val Thr Asn Val Ser Gln Ala Pro Val Tyr Ile Ser
            340                 345                 350

acc ggg acg cag ctg ggg caa gcc atc ttc gtg ttc gcg ccg cgg ttc     1104
Thr Gly Thr Gln Leu Gly Gln Ala Ile Phe Val Phe Ala Pro Arg Phe
        355                 360                 365

ggt ggc ccg gcg aaa ctg cgg cag ctc ctc ggc cac cga tcg cgc gcc     1152
Gly Gly Pro Ala Lys Leu Arg Gln Leu Leu Gly His Arg Ser Arg Ala
    370                 375                 380

ctg gag ctg ccg ggc ggg gtg aca gtg gac agc caa aaa ctg tgt agg     1200
Leu Glu Leu Pro Gly Gly Val Thr Val Asp Ser Gln Lys Leu Cys Arg
385                 390                 395                 400

ttt gag acc atg tac ctg ttt tcc acg taa                             1230
Phe Glu Thr Met Tyr Leu Phe Ser Thr
                405                 410


<210> 19
<211> 409
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 19
Met Gly Thr Pro Val Arg Phe Phe Arg Gly Glu Trp Gln Thr Ser Ser
 1               5                  10                  15
Leu Val Asp Asn Gly Thr Pro Arg Tyr Ser Ser Leu Val Trp Ala Ala
            20                  25                  30
Thr Ile His Asp Gly Tyr Leu Thr Leu Val Asn Arg Ser Glu Leu Cys
        35                  40                  45
Val Thr Glu Arg Ser Pro Cys Leu Pro Ala Cys Pro Ser Ile Gly Arg
    50                  55                  60
Leu Val Gly Lys Arg Phe Pro Gly Phe Ala Phe Ala Ser Ala Thr Leu
65                  70                  75                  80
Gly Asp Arg Gly Thr Arg Thr Val Phe Tyr Ala Phe Gly His Arg Asp
                85                  90                  95
Asn Pro Leu Asp Ile Val Pro Ala Val Val Glu Arg Ala Asp Arg Glu
            100                 105                 110
Leu Val Leu Arg Val His Ala Pro Gln Thr Thr Arg Val Ser Arg Tyr
        115                 120                 125
Gly Leu Lys Val Phe Val Ala Ile Val Thr Val Val Arg Pro Pro Gly
    130                 135                 140
```

Val Phe Leu His Phe Pro Gln Asp Arg Val Pro Ile Ala Leu Thr Asp
145 150 155 160
Ala Cys Ser Gln Glu Gly Ser Arg Leu Thr Ser Glu Glu Pro Trp Ile
165 170 175
Lys Ile Gln Gly Phe Pro Val Leu Ser Asp Glu Thr Ala His Pro Phe
180 185 190
Leu Leu Thr Gln Lys Thr Lys Pro Phe Thr Glu Arg Lys Phe Cys Arg
195 200 205
Leu Ile Met Asp Asn Asp Gln Arg Ser Ala Val Asn Thr Val Tyr Leu
210 215 220
Gly Lys Gln His Val Arg Val Thr Val Thr Arg Pro Pro Glu Thr Ile
225 230 235 240
Val Thr Asp Gly Pro Val Thr Ala Thr Leu Ser Leu Thr Gly Asn Ala
245 250 255
Pro Ile Ala Phe Arg His Asn Pro Tyr Phe Glu Leu Pro Trp Ser Ser
260 265 270
Thr Thr Ala Ile Phe Thr Pro Val Val Tyr Val Gly Leu Thr Val Cys
275 280 285
Ile Pro Pro Asn Cys Ser Lys Phe Val Arg Tyr Gly Asn Thr Tyr Val
290 295 300
Ser Ala Phe Asn Arg Lys Leu Thr Ala Ile Ile Ser Asn His Ala His
305 310 315 320
Asn Gly Gly Phe Arg Ile Gln Asp Cys Glu Trp Pro Pro Asn Arg Glu
325 330 335
Ile Glu Ile Leu Val Thr Asn Val Ser Gln Ala Pro Val Tyr Ile Ser
340 345 350
Thr Gly Thr Gln Leu Gly Gln Ala Ile Phe Val Phe Ala Pro Arg Phe
355 360 365
Gly Gly Pro Ala Lys Leu Arg Gln Leu Leu Gly His Arg Ser Arg Ala
370 375 380
Leu Glu Leu Pro Gly Gly Val Thr Val Asp Ser Gln Lys Leu Cys Arg
385 390 395 400
Phe Glu Thr Met Tyr Leu Phe Ser Thr
405

<210> 20
<211> 624
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(624)

<400> 20
atg ttc cct gtc tgg ttc gtc ttg ttt tac ctg tcg tgt tgg gcg gcc 48
Met Phe Pro Val Trp Phe Val Leu Phe Tyr Leu Ser Cys Trp Ala Ala
1 5 10 15

agc cct acg ctg gcg cct ccc ccg act gcc gct gga att aac gtt ctc 96
Ser Pro Thr Leu Ala Pro Pro Pro Thr Ala Ala Gly Ile Asn Val Leu
20 25 30

ccc cag tgg gcc ggc aac cgc gcc tct ctt gac agg acc agg ggg cgc 144
Pro Gln Trp Ala Gly Asn Arg Ala Ser Leu Asp Arg Thr Arg Gly Arg
35 40 45

ctg tct gaa gtg ggg tta aac ata cag cgc tgg ttc gtt tac ctg tgc 192
Leu Ser Glu Val Gly Leu Asn Ile Gln Arg Trp Phe Val Tyr Leu Cys

```
     50                  55                  60

cac cac tcc act ctc tgt cgg gtg cgt gag tac ccg cgc atc atg tcg     240
His His Ser Thr Leu Cys Arg Val Arg Glu Tyr Pro Arg Ile Met Ser
 65                 70                  75                  80

ttt gtt cac ttc cct ata ttg atg tct aac gtt gag tgc cag cgc cgc     288
Phe Val His Phe Pro Ile Leu Met Ser Asn Val Glu Cys Gln Arg Arg
                85                  90                  95

gag ttt cgc ggg gcc gag tgt atg aac gcc atg gtt cgc ggg ctc cgg     336
Glu Phe Arg Gly Ala Glu Cys Met Asn Ala Met Val Arg Gly Leu Arg
            100                 105                 110

gcc tac gag agt tac ctg acg cga ctg agg atg ctg ctg gac gac gcg     384
Ala Tyr Glu Ser Tyr Leu Thr Arg Leu Arg Met Leu Leu Asp Asp Ala
        115                 120                 125

ccc ggg gac gcg gac gcc gcg gcc att ggc tcc gcg gtg acc gtg gtg     432
Pro Gly Asp Ala Asp Ala Ala Ala Ile Gly Ser Ala Val Thr Val Val
    130                 135                 140

ctg tcc gcc ctc gac tct cta att gag gag ctt ccc gta aat aac aag     480
Leu Ser Ala Leu Asp Ser Leu Ile Glu Glu Leu Pro Val Asn Asn Lys
145                 150                 155                 160

ata ggt ggc gcg gag tct aat gaa aaa acc gtg cgt gcg ttg gga ggg     528
Ile Gly Gly Ala Glu Ser Asn Glu Lys Thr Val Arg Ala Leu Gly Gly
                165                 170                 175

cag agc ccc cgg gac gtt gtt ctc agc gcg ttt cgc ata ctg gaa tat     576
Gln Ser Pro Arg Asp Val Val Leu Ser Ala Phe Arg Ile Leu Glu Tyr
            180                 185                 190

cta cag atg ttt ttg cgg gac ggg cgc cgc gca ata gct atg atg taa     624
Leu Gln Met Phe Leu Arg Asp Gly Arg Arg Ala Ile Ala Met Met
        195                 200                 205


<210> 21
<211> 207
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 21
Met Phe Pro Val Trp Phe Val Leu Phe Tyr Leu Ser Cys Trp Ala Ala
  1               5                  10                  15
Ser Pro Thr Leu Ala Pro Pro Pro Thr Ala Ala Gly Ile Asn Val Leu
             20                  25                  30
Pro Gln Trp Ala Gly Asn Arg Ala Ser Leu Asp Arg Thr Arg Gly Arg
         35                  40                  45
Leu Ser Glu Val Gly Leu Asn Ile Gln Arg Trp Phe Val Tyr Leu Cys
     50                  55                  60
His His Ser Thr Leu Cys Arg Val Arg Glu Tyr Pro Arg Ile Met Ser
 65                  70                  75                  80
Phe Val His Phe Pro Ile Leu Met Ser Asn Val Glu Cys Gln Arg Arg
                 85                  90                  95
Glu Phe Arg Gly Ala Glu Cys Met Asn Ala Met Val Arg Gly Leu Arg
            100                 105                 110
Ala Tyr Glu Ser Tyr Leu Thr Arg Leu Arg Met Leu Leu Asp Asp Ala
        115                 120                 125
```

Pro Gly Asp Ala Asp Ala Ala Ala Ile Gly Ser Ala Val Thr Val Val
130 135 140
Leu Ser Ala Leu Asp Ser Leu Ile Glu Glu Leu Pro Val Asn Asn Lys
145 150 155 160
Ile Gly Gly Ala Glu Ser Asn Glu Lys Thr Val Arg Ala Leu Gly Gly
165 170 175
Gln Ser Pro Arg Asp Val Val Leu Ser Ala Phe Arg Ile Leu Glu Tyr
180 185 190
Leu Gln Met Phe Leu Arg Asp Gly Arg Arg Ala Ile Ala Met Met
195 200 205

<210> 22
<211> 1002
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1002)

<400> 22
atg att gtc ctg gtg cat ctc ggt atc tgc tac gtt aaa aaa att ata 48
Met Ile Val Leu Val His Leu Gly Ile Cys Tyr Val Lys Lys Ile Ile
1 5 10 15

ccc gtt tgc gtg gct gga atc gcg gcc gca cga ctc cgg gtt ttt tcc 96
Pro Val Cys Val Ala Gly Ile Ala Ala Ala Arg Leu Arg Val Phe Ser
20 25 30

gca ccc gaa ggt gcc gcg gcc gtc cgc tgc gcg tgc cgc ggc gac cac 144
Ala Pro Glu Gly Ala Ala Ala Val Arg Cys Ala Cys Arg Gly Asp His
35 40 45

ggc gag ctg cag tac ttg gcg cac ctg gat tta att att aaa cac ggc 192
Gly Glu Leu Gln Tyr Leu Ala His Leu Asp Leu Ile Ile Lys His Gly
50 55 60

gtg cag agg gag gac agg acg ggc gtg ggc acc agg tcc gtg ttc ggg 240
Val Gln Arg Glu Asp Arg Thr Gly Val Gly Thr Arg Ser Val Phe Gly
65 70 75 80

ctt cag gcc cga tat aac ctc agg gac gag ttt cct ctg tta acc acc 288
Leu Gln Ala Arg Tyr Asn Leu Arg Asp Glu Phe Pro Leu Leu Thr Thr
85 90 95

aaa agg gtg ttt tgg agg ggc gtc gtg gag gag ttg ctg tgg ttt atc 336
Lys Arg Val Phe Trp Arg Gly Val Val Glu Glu Leu Leu Trp Phe Ile
100 105 110

agg ggc tcc acc gac tcc acc gaa ctg tcg cgc cgc ggc gta aaa att 384
Arg Gly Ser Thr Asp Ser Thr Glu Leu Ser Arg Arg Gly Val Lys Ile
115 120 125

tgg gac gcg cac ggg tcc cgc gcc ttt ttg gcg gcg cag ggc ttc ggg 432
Trp Asp Ala His Gly Ser Arg Ala Phe Leu Ala Ala Gln Gly Phe Gly
130 135 140

gac cgc cgc gag ggc gat ctg ggg ccg gtg tac ggg ttc cag tgg aga 480
Asp Arg Arg Glu Gly Asp Leu Gly Pro Val Tyr Gly Phe Gln Trp Arg

```
145                 150                 155                 160

cat ttt ggg gcg gag tac agg ggg gcc gac gcc aac tac gag ggt cag      528
His Phe Gly Ala Glu Tyr Arg Gly Ala Asp Ala Asn Tyr Glu Gly Gln
                165                 170                 175

ggg gtg gac cag ctg cgt tac gtg gtg gat cta att aac agg cgg ccc      576
Gly Val Asp Gln Leu Arg Tyr Val Val Asp Leu Ile Asn Arg Arg Pro
            180                 185                 190

cac gat cgg cgc atc gtt atg tgc gcg tgg aac ccc gcg gac ctc gcg      624
His Asp Arg Arg Ile Val Met Cys Ala Trp Asn Pro Ala Asp Leu Ala
        195                 200                 205

cgg atg gct ctc cct cct tgt cac gtt ttg tgt cag ttt tac gtg gct      672
Arg Met Ala Leu Pro Pro Cys His Val Leu Cys Gln Phe Tyr Val Ala
    210                 215                 220

cgg ggg gag ctg tcc tgc cag ctg tac cag agg tcc gcc gac atg ggc      720
Arg Gly Glu Leu Ser Cys Gln Leu Tyr Gln Arg Ser Ala Asp Met Gly
225                 230                 235                 240

ctc ggg gtc ccg ttt aac atc gcc agc tac gcc ctc ctg acg tat ctg      768
Leu Gly Val Pro Phe Asn Ile Ala Ser Tyr Ala Leu Leu Thr Tyr Leu
                245                 250                 255

atc gct cac gtc acg ggc ctg acc ccg gga gac ttt gtg cac acc ttg      816
Ile Ala His Val Thr Gly Leu Thr Pro Gly Asp Phe Val His Thr Leu
            260                 265                 270

ggt gac gcc cac gtt tac aac aac cac gtt gat ccc ctg ctg ctt cag      864
Gly Asp Ala His Val Tyr Asn Asn His Val Asp Pro Leu Leu Leu Gln
        275                 280                 285

ctg cgg agg acc ccg cgt ccg ttt ccg cgg ctg aag att ttg aga aag      912
Leu Arg Arg Thr Pro Arg Pro Phe Pro Arg Leu Lys Ile Leu Arg Lys
    290                 295                 300

gtg gcg cgt ctg gag gac ttt acg cgc gcg gat ctg agt ctc gag ggc      960
Val Ala Arg Leu Glu Asp Phe Thr Arg Ala Asp Leu Ser Leu Glu Gly
305                 310                 315                 320

tac gac ccc cat ccc cac ata gag atg gag atg gcc gtt tga              1002
Tyr Asp Pro His Pro His Ile Glu Met Glu Met Ala Val
                325                 330


<210> 23
<211> 333
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 23
Met Ile Val Leu Val His Leu Gly Ile Cys Tyr Val Lys Lys Ile Ile
  1               5                  10                  15
Pro Val Cys Val Ala Gly Ile Ala Ala Ala Arg Leu Arg Val Phe Ser
            20                  25                  30
Ala Pro Glu Gly Ala Ala Ala Val Arg Cys Ala Cys Arg Gly Asp His
        35                  40                  45
Gly Glu Leu Gln Tyr Leu Ala His Leu Asp Leu Ile Ile Lys His Gly
    50                  55                  60
```

Val Gln Arg Glu Asp Arg Thr Gly Val Gly Thr Arg Ser Val Phe Gly
65 70 75 80
Leu Gln Ala Arg Tyr Asn Leu Arg Asp Glu Phe Pro Leu Leu Thr Thr
85 90 95
Lys Arg Val Phe Trp Arg Gly Val Val Glu Glu Leu Leu Trp Phe Ile
100 105 110
Arg Gly Ser Thr Asp Ser Thr Glu Leu Ser Arg Arg Gly Val Lys Ile
115 120 125
Trp Asp Ala His Gly Ser Arg Ala Phe Leu Ala Ala Gln Gly Phe Gly
130 135 140
Asp Arg Arg Glu Gly Asp Leu Gly Pro Val Tyr Gly Phe Gln Trp Arg
145 150 155 160
His Phe Gly Ala Glu Tyr Arg Gly Ala Asp Ala Asn Tyr Glu Gly Gln
165 170 175
Gly Val Asp Gln Leu Arg Tyr Val Val Asp Leu Ile Asn Arg Arg Pro
180 185 190
His Asp Arg Arg Ile Val Met Cys Ala Trp Asn Pro Ala Asp Leu Ala
195 200 205
Arg Met Ala Leu Pro Pro Cys His Val Leu Cys Gln Phe Tyr Val Ala
210 215 220
Arg Gly Glu Leu Ser Cys Gln Leu Tyr Gln Arg Ser Ala Asp Met Gly
225 230 235 240
Leu Gly Val Pro Phe Asn Ile Ala Ser Tyr Ala Leu Leu Thr Tyr Leu
245 250 255
Ile Ala His Val Thr Gly Leu Thr Pro Gly Asp Phe Val His Thr Leu
260 265 270
Gly Asp Ala His Val Tyr Asn Asn His Val Asp Pro Leu Leu Leu Gln
275 280 285
Leu Arg Arg Thr Pro Arg Pro Phe Pro Arg Leu Lys Ile Leu Arg Lys
290 295 300
Val Ala Arg Leu Glu Asp Phe Thr Arg Ala Asp Leu Ser Leu Glu Gly
305 310 315 320
Tyr Asp Pro His Pro His Ile Glu Met Glu Met Ala Val
325 330

<210> 24
<211> 348
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(348)

<400> 24
atg agg ggc ctt ttc gtg tgc gtt ttt ttt gcg gtg ttc gcg tgt gta 48
Met Arg Gly Leu Phe Val Cys Val Phe Phe Ala Val Phe Ala Cys Val
1 5 10 15

gtt gat tat gcc ttt cct atg ggc tcc atg agc gga ccc gcg ccc gaa 96
Val Asp Tyr Ala Phe Pro Met Gly Ser Met Ser Gly Pro Ala Pro Glu
20 25 30

ctc tgc tgt ttg ggg tat gta act cat ctg ccg cca ccc ggt tta gtg 144
Leu Cys Cys Leu Gly Tyr Val Thr His Leu Pro Pro Pro Gly Leu Val
35 40 45

gtc tct tac tcc cac acc tcg tcg cag tgc tcg gtg gac gcc gtg ata 192
Val Ser Tyr Ser His Thr Ser Ser Gln Cys Ser Val Asp Ala Val Ile

50 55 60

tta aac act cgc cgc ggt aaa aag ctg tgt gcc aat ccc ggg gac gac 240
Leu Asn Thr Arg Arg Gly Lys Lys Leu Cys Ala Asn Pro Gly Asp Asp
65 70 75 80

gca gtg aag aaa ctg ctt cag gcg gtg gac aag cgt ccc aaa aag ggc 288
Ala Val Lys Lys Leu Leu Gln Ala Val Asp Lys Arg Pro Lys Lys Gly
85 90 95

aga aga acc cgg cgc agc ctg att gac gat tcc gaa gag ggc ctt ggc 336
Arg Arg Thr Arg Arg Ser Leu Ile Asp Asp Ser Glu Glu Gly Leu Gly
100 105 110

agc ggg att tag 348
Ser Gly Ile
115

<210> 25
<211> 115
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 25
Met Arg Gly Leu Phe Val Cys Val Phe Phe Ala Val Phe Ala Cys Val
1 5 10 15
Val Asp Tyr Ala Phe Pro Met Gly Ser Met Ser Gly Pro Ala Pro Glu
20 25 30
Leu Cys Cys Leu Gly Tyr Val Thr His Leu Pro Pro Pro Gly Leu Val
35 40 45
Val Ser Tyr Ser His Thr Ser Ser Gln Cys Ser Val Asp Ala Val Ile
50 55 60
Leu Asn Thr Arg Arg Gly Lys Lys Leu Cys Ala Asn Pro Gly Asp Asp
65 70 75 80
Ala Val Lys Lys Leu Leu Gln Ala Val Asp Lys Arg Pro Lys Lys Gly
85 90 95
Arg Arg Thr Arg Arg Ser Leu Ile Asp Asp Ser Glu Glu Gly Leu Gly
100 105 110
Ser Gly Ile
115

<210> 26
<211> 564
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(564)

<400> 26
atg gcg gct gtt cag ggc cct ccg ccg ccc cct gag gaa gaa aat gaa 48
Met Ala Ala Val Gln Gly Pro Pro Pro Pro Pro Glu Glu Glu Asn Glu
1 5 10 15

aat tct ttg cca gtt gac gtt tat gct ata gag ggc atc ttt ctt tac 96
Asn Ser Leu Pro Val Asp Val Tyr Ala Ile Glu Gly Ile Phe Leu Tyr
20 25 30

tgt ggg ctc ggg cag gcg gag tac ttg cac cat ccc gtg ttt agc cct 144
Cys Gly Leu Gly Gln Ala Glu Tyr Leu His His Pro Val Phe Ser Pro
35 40 45

att aag gaa ttt atc agc gcc ttt ctc aag gac agc gct cgc ctg tac 192
Ile Lys Glu Phe Ile Ser Ala Phe Leu Lys Asp Ser Ala Arg Leu Tyr
50 55 60

gag agg ctc ttg cgc cac acc gat tac cgc tct ctg cgg gga cta aac 240
Glu Arg Leu Leu Arg His Thr Asp Tyr Arg Ser Leu Arg Gly Leu Asn
65 70 75 80

gcc ata ggc caa ggg atg ctg caa ata aac acg gac gga cgc cac aac 288
Ala Ile Gly Gln Gly Met Leu Gln Ile Asn Thr Asp Gly Arg His Asn
85 90 95

tgg ggt cgc gct ttg gcc gtg ttg ggt ctt ggt gcg tat gtg gtg gat 336
Trp Gly Arg Ala Leu Ala Val Leu Gly Leu Gly Ala Tyr Val Val Asp
100 105 110

aag gtt aaa gac gac gag cgt ctt tta acg ttc gcc ata gcc gtt cta 384
Lys Val Lys Asp Asp Glu Arg Leu Leu Thr Phe Ala Ile Ala Val Leu
115 120 125

ccc gtg tac gcg tac gag gcg ctg gag tct cag tgg ttt cgt tca cac 432
Pro Val Tyr Ala Tyr Glu Ala Leu Glu Ser Gln Trp Phe Arg Ser His
130 135 140

ggc gaa tgg gag gga ctc agg aat tac tgc gag cga ata ctg agg cat 480
Gly Glu Trp Glu Gly Leu Arg Asn Tyr Cys Glu Arg Ile Leu Arg His
145 150 155 160

cgc cgc aac gcg agg aga cac atg tgc tac gga gtt gcg gct ggt ctt 528
Arg Arg Asn Ala Arg Arg His Met Cys Tyr Gly Val Ala Ala Gly Leu
165 170 175

ctg gcg cta gtg gcg ctg ttt gcc atc agg cga tag 564
Leu Ala Leu Val Ala Leu Phe Ala Ile Arg Arg
180 185

<210> 27
<211> 187
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 27
Met Ala Ala Val Gln Gly Pro Pro Pro Pro Pro Glu Glu Glu Asn Glu
1 5 10 15
Asn Ser Leu Pro Val Asp Val Tyr Ala Ile Glu Gly Ile Phe Leu Tyr
20 25 30
Cys Gly Leu Gly Gln Ala Glu Tyr Leu His His Pro Val Phe Ser Pro
35 40 45
Ile Lys Glu Phe Ile Ser Ala Phe Leu Lys Asp Ser Ala Arg Leu Tyr
50 55 60
Glu Arg Leu Leu Arg His Thr Asp Tyr Arg Ser Leu Arg Gly Leu Asn
65 70 75 80
Ala Ile Gly Gln Gly Met Leu Gln Ile Asn Thr Asp Gly Arg His Asn
85 90 95
Trp Gly Arg Ala Leu Ala Val Leu Gly Leu Gly Ala Tyr Val Val Asp

```
            100                     105                     110
Lys Val Lys Asp Asp Glu Arg Leu Leu Thr Phe Ala Ile Ala Val Leu
        115                     120                     125
Pro Val Tyr Ala Tyr Glu Ala Leu Glu Ser Gln Trp Phe Arg Ser His
    130                     135                     140
Gly Glu Trp Glu Gly Leu Arg Asn Tyr Cys Glu Arg Ile Leu Arg His
145                     150                     155                 160
Arg Arg Asn Ala Arg Arg His Met Cys Tyr Gly Val Ala Ala Gly Leu
                165                     170                     175
Leu Ala Leu Val Ala Leu Phe Ala Ile Arg Arg
            180                     185
```

<210> 28
<211> 1611
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1611)

<400> 28

```
atg act ccc gtg tac gtt ggg gga tac gtg gac gtg gtc agc cta cca      48
Met Thr Pro Val Tyr Val Gly Gly Tyr Val Asp Val Val Ser Leu Pro
1               5                   10                  15

aag ata gaa aag gag ctg tat tta gag ccc tca atc gtg gcg acc ctg      96
Lys Ile Glu Lys Glu Leu Tyr Leu Glu Pro Ser Ile Val Ala Thr Leu
            20                  25                  30

ctc ccg tat acg gac cct cta ccg ata aac ata gag cac gtc ccc gaa     144
Leu Pro Tyr Thr Asp Pro Leu Pro Ile Asn Ile Glu His Val Pro Glu
        35                  40                  45

gcc cac gta ggt cac aca atc ggt ctc ttc caa gta aca cac ggg ata     192
Ala His Val Gly His Thr Ile Gly Leu Phe Gln Val Thr His Gly Ile
    50                  55                  60

ttt tgc ctg ggc aag cta acg agc cac gat ttc ttg gcc ctg gcg tca     240
Phe Cys Leu Gly Lys Leu Thr Ser His Asp Phe Leu Ala Leu Ala Ser
65                  70                  75                  80

cgg ctc gcg ggg gac tcg cga gcg gcc cag atc cag cta aac ccc atg     288
Arg Leu Ala Gly Asp Ser Arg Ala Ala Gln Ile Gln Leu Asn Pro Met
                85                  90                  95

cct cgg gac ccc ctg ttg gag atg tta cac aca tgg ctc ccg gag ctc     336
Pro Arg Asp Pro Leu Leu Glu Met Leu His Thr Trp Leu Pro Glu Leu
            100                 105                 110

tcc ctc tcg tcc ctc cac cca gaa gag ctg caa gac ccc aat cac cca     384
Ser Leu Ser Ser Leu His Pro Glu Glu Leu Gln Asp Pro Asn His Pro
        115                 120                 125

ccc gcg ttt cag cac gtc tct ctg tgc gca ctg ggg agg cgt agg ggg     432
Pro Ala Phe Gln His Val Ser Leu Cys Ala Leu Gly Arg Arg Arg Gly
    130                 135                 140

tca atc gcc gtt tac gga cca gac cct aca tgg gtc gtc tcc aag ttt     480
```

Ser Ile Ala Val Tyr Gly Pro Asp Pro Thr Trp Val Val Ser Lys Phe
145 150 155 160

gat tct ctt acg cga gag gaa gcc ggt aaa ata acg gtc aat tgt tta 528
Asp Ser Leu Thr Arg Glu Glu Ala Gly Lys Ile Thr Val Asn Cys Leu
165 170 175

gat ctt tgc gaa cgt cag gtg aca ccg cca gaa ttc gcg gcc cct ctt 576
Asp Leu Cys Glu Arg Gln Val Thr Pro Pro Glu Phe Ala Ala Pro Leu
180 185 190

gag aca ctg atg gcc aag gcg atc gac gcc gga ttt atc cga gat cgc 624
Glu Thr Leu Met Ala Lys Ala Ile Asp Ala Gly Phe Ile Arg Asp Arg
195 200 205

aca gac ctg ctc aaa aca gat aaa ggc gtg gcc aga gtc gcg aga agt 672
Thr Asp Leu Leu Lys Thr Asp Lys Gly Val Ala Arg Val Ala Arg Ser
210 215 220

acg tat tta aag gca agc cag ttt ccc tgc gct cag cac tgc ggt aac 720
Thr Tyr Leu Lys Ala Ser Gln Phe Pro Cys Ala Gln His Cys Gly Asn
225 230 235 240

cgc gac acc cga acc atg agc gcc ctc ccg gag gac aac atc acc att 768
Arg Asp Thr Arg Thr Met Ser Ala Leu Pro Glu Asp Asn Ile Thr Ile
245 250 255

ccc aag agc acc ttt ctg acc atg gtg cag agc agt ctc gat aac atg 816
Pro Lys Ser Thr Phe Leu Thr Met Val Gln Ser Ser Leu Asp Asn Met
260 265 270

cgc aac cag ggc cac cgc acg tac gtt tct gcg cca ccc tcg atg ccg 864
Arg Asn Gln Gly His Arg Thr Tyr Val Ser Ala Pro Pro Ser Met Pro
275 280 285

gca acg gcg gcg tat cca tcg tgg ata ccg ccg cct gag ctg acc gtc 912
Ala Thr Ala Ala Tyr Pro Ser Trp Ile Pro Pro Pro Glu Leu Thr Val
290 295 300

cca tcg tac gca ccg ccc gtc gcg cca ccg ttc cct ttc cag tcg gcg 960
Pro Ser Tyr Ala Pro Pro Val Ala Pro Pro Phe Pro Phe Gln Ser Ala
305 310 315 320

ttt gcc ccc caa ccc tct ccg tat gcg gcc acg tac tat tct ccg acg 1008
Phe Ala Pro Gln Pro Ser Pro Tyr Ala Ala Thr Tyr Tyr Ser Pro Thr
325 330 335

tac ggc tac gct ccg gca ccg tct cgc cac cag aag cgc aaa cgg gac 1056
Tyr Gly Tyr Ala Pro Ala Pro Ser Arg His Gln Lys Arg Lys Arg Asp
340 345 350

gtg gag ctt tcc gac gag ccc gtg ttt cca ggc gag gag gtc ggc ata 1104
Val Glu Leu Ser Asp Glu Pro Val Phe Pro Gly Glu Glu Val Gly Ile
355 360 365

cac aag gac gtc atg gcg ctg tcc aaa aac ata ttg gac att cag gcg 1152
His Lys Asp Val Met Ala Leu Ser Lys Asn Ile Leu Asp Ile Gln Ala
370 375 380

gat ctg agg gac cta aaa cga gcc gcg tcg cag acc tcc ggc gca cag 1200
Asp Leu Arg Asp Leu Lys Arg Ala Ala Ser Gln Thr Ser Gly Ala Gln

385 390 395 400

gac gcc gac caa cga ccg caa cca ccg ccc gtt cag ttc tcg tgg cca 1248
Asp Ala Asp Gln Arg Pro Gln Pro Pro Pro Val Gln Phe Ser Trp Pro
405 410 415

cag acc tac gcg tct gcg ccg tac cta gcg tac cag ccg cag tgg tat 1296
Gln Thr Tyr Ala Ser Ala Pro Tyr Leu Ala Tyr Gln Pro Gln Trp Tyr
420 425 430

tcc gga acg gac acc cat ctc cac gcc cca cag ccg tac cag agc gcg 1344
Ser Gly Thr Asp Thr His Leu His Ala Pro Gln Pro Tyr Gln Ser Ala
435 440 445

cag ggc atc cag caa acg caa cca ccg ccc ccg cag ccg gcg agc cac 1392
Gln Gly Ile Gln Gln Thr Gln Pro Pro Pro Pro Gln Pro Ala Ser His
450 455 460

cac gcc ggt ctc gcc acg cag ccg gca act cca gcc ccc gcc gcc caa 1440
His Ala Gly Leu Ala Thr Gln Pro Ala Thr Pro Ala Pro Ala Ala Gln
465 470 475 480

gaa tct gta atg tca aac gcc att cct tca gcc tct gcc ccc cgc gct 1488
Glu Ser Val Met Ser Asn Ala Ile Pro Ser Ala Ser Ala Pro Arg Ala
485 490 495

ggc gcg tgc ccg ccc ctc gat cca gaa tgc gga cag tcc gcg cgg gcc 1536
Gly Ala Cys Pro Pro Leu Asp Pro Glu Cys Gly Gln Ser Ala Arg Ala
500 505 510

ccg gtg gag gcc agc gca cag ccg gcc ccc gtg tca cag ata caa aaa 1584
Pro Val Glu Ala Ser Ala Gln Pro Ala Pro Val Ser Gln Ile Gln Lys
515 520 525

atg ttc tgc gag gaa ctg ctc aaa taa 1611
Met Phe Cys Glu Glu Leu Leu Lys
530 535

<210> 29
<211> 536
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 29
Met Thr Pro Val Tyr Val Gly Gly Tyr Val Asp Val Val Ser Leu Pro
1 5 10 15
Lys Ile Glu Lys Glu Leu Tyr Leu Glu Pro Ser Ile Val Ala Thr Leu
20 25 30
Leu Pro Tyr Thr Asp Pro Leu Pro Ile Asn Ile Glu His Val Pro Glu
35 40 45
Ala His Val Gly His Thr Ile Gly Leu Phe Gln Val Thr His Gly Ile
50 55 60
Phe Cys Leu Gly Lys Leu Thr Ser His Asp Phe Leu Ala Leu Ala Ser
65 70 75 80
Arg Leu Ala Gly Asp Ser Arg Ala Ala Gln Ile Gln Leu Asn Pro Met
85 90 95
Pro Arg Asp Pro Leu Leu Glu Met Leu His Thr Trp Leu Pro Glu Leu
100 105 110
Ser Leu Ser Ser Leu His Pro Glu Glu Leu Gln Asp Pro Asn His Pro
115 120 125

Pro Ala Phe Gln His Val Ser Leu Cys Ala Leu Gly Arg Arg Arg Gly
130 135 140
Ser Ile Ala Val Tyr Gly Pro Asp Pro Thr Trp Val Val Ser Lys Phe
145 150 155 160
Asp Ser Leu Thr Arg Glu Glu Ala Gly Lys Ile Thr Val Asn Cys Leu
165 170 175
Asp Leu Cys Glu Arg Gln Val Thr Pro Pro Glu Phe Ala Ala Pro Leu
180 185 190
Glu Thr Leu Met Ala Lys Ala Ile Asp Ala Gly Phe Ile Arg Asp Arg
195 200 205
Thr Asp Leu Leu Lys Thr Asp Lys Gly Val Ala Arg Val Ala Arg Ser
210 215 220
Thr Tyr Leu Lys Ala Ser Gln Phe Pro Cys Ala Gln His Cys Gly Asn
225 230 235 240
Arg Asp Thr Arg Thr Met Ser Ala Leu Pro Glu Asp Asn Ile Thr Ile
245 250 255
Pro Lys Ser Thr Phe Leu Thr Met Val Gln Ser Ser Leu Asp Asn Met
260 265 270
Arg Asn Gln Gly His Arg Thr Tyr Val Ser Ala Pro Pro Ser Met Pro
275 280 285
Ala Thr Ala Ala Tyr Pro Ser Trp Ile Pro Pro Pro Glu Leu Thr Val
290 295 300
Pro Ser Tyr Ala Pro Pro Val Ala Pro Pro Phe Pro Phe Gln Ser Ala
305 310 315 320
Phe Ala Pro Gln Pro Ser Pro Tyr Ala Ala Thr Tyr Tyr Ser Pro Thr
325 330 335
Tyr Gly Tyr Ala Pro Ala Pro Ser Arg His Gln Lys Arg Lys Arg Asp
340 345 350
Val Glu Leu Ser Asp Glu Pro Val Phe Pro Gly Glu Glu Val Gly Ile
355 360 365
His Lys Asp Val Met Ala Leu Ser Lys Asn Ile Leu Asp Ile Gln Ala
370 375 380
Asp Leu Arg Asp Leu Lys Arg Ala Ala Ser Gln Thr Ser Gly Ala Gln
385 390 395 400
Asp Ala Asp Gln Arg Pro Gln Pro Pro Pro Val Gln Phe Ser Trp Pro
405 410 415
Gln Thr Tyr Ala Ser Ala Pro Tyr Leu Ala Tyr Gln Pro Gln Trp Tyr
420 425 430
Ser Gly Thr Asp Thr His Leu His Ala Pro Gln Pro Tyr Gln Ser Ala
435 440 445
Gln Gly Ile Gln Gln Thr Gln Pro Pro Pro Pro Gln Pro Ala Ser His
450 455 460
His Ala Gly Leu Ala Thr Gln Pro Ala Thr Pro Ala Pro Ala Ala Gln
465 470 475 480
Glu Ser Val Met Ser Asn Ala Ile Pro Ser Ala Ser Ala Pro Arg Ala
485 490 495
Gly Ala Cys Pro Pro Leu Asp Pro Glu Cys Gly Gln Ser Ala Arg Ala
500 505 510
Pro Val Glu Ala Ser Ala Gln Pro Ala Pro Val Ser Gln Ile Gln Lys
515 520 525
Met Phe Cys Glu Glu Leu Leu Lys
530 535

<210> 30
<211> 900
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>

<221> CDS
<222> (1)..(900)

<400> 30
atg ttt atc ggt aga ggg tcc gta tac ggg agc agg gtc gcc acg att 48
Met Phe Ile Gly Arg Gly Ser Val Tyr Gly Ser Arg Val Ala Thr Ile
1 5 10 15

gag ggc tct aaa tac agc tcc ttt tct atc ttt ggt agg ctg acc acg 96
Glu Gly Ser Lys Tyr Ser Ser Phe Ser Ile Phe Gly Arg Leu Thr Thr
20 25 30

tcc acg tat ccc cca acg tac acg gga gtc atg ttg ggt cgg tgt ctg 144
Ser Thr Tyr Pro Pro Thr Tyr Thr Gly Val Met Leu Gly Arg Cys Leu
35 40 45

cgt gag ccg aag gag atg agc gcc ggt ctt cgg ggg ctg atg tgg cgc 192
Arg Glu Pro Lys Glu Met Ser Ala Gly Leu Arg Gly Leu Met Trp Arg
50 55 60

gtt atc cgt tgt gag aat tta aat acg ttt ctc cct ggc gag ctc agg 240
Val Ile Arg Cys Glu Asn Leu Asn Thr Phe Leu Pro Gly Glu Leu Arg
65 70 75 80

ttt ctt cac ctg gtg ctg tgc gaa atg tac aac tac ggt ctt aac gtg 288
Phe Leu His Leu Val Leu Cys Glu Met Tyr Asn Tyr Gly Leu Asn Val
85 90 95

tac ctt ctg aag gag gca att gcc aat acc gga acg agg gac gat att 336
Tyr Leu Leu Lys Glu Ala Ile Ala Asn Thr Gly Thr Arg Asp Asp Ile
100 105 110

gtc ctc ggg cgt aag gtt ccg gtg gag ttc tgg aaa atc gtt tac gat 384
Val Leu Gly Arg Lys Val Pro Val Glu Phe Trp Lys Ile Val Tyr Asp
115 120 125

ggg tta aag gag atg ggg gtt tct gac gca acg ttg ctc tca gaa acc 432
Gly Leu Lys Glu Met Gly Val Ser Asp Ala Thr Leu Leu Ser Glu Thr
130 135 140

aaa agg ggg gct ctg tgg ctt tac ttc aac ggg cgt ccg tgc ctg tta 480
Lys Arg Gly Ala Leu Trp Leu Tyr Phe Asn Gly Arg Pro Cys Leu Leu
145 150 155 160

aaa ggt cta ggc gac tac gtt ttt tgc caa ctg ggg tta agc cac agc 528
Lys Gly Leu Gly Asp Tyr Val Phe Cys Gln Leu Gly Leu Ser His Ser
165 170 175

gta cgc gtg gtt ccg gaa aac cta acg gat ggg aat tat ctc tat aac 576
Val Arg Val Val Pro Glu Asn Leu Thr Asp Gly Asn Tyr Leu Tyr Asn
180 185 190

cta ggc agc gtc ata cca tgt cgc ctg ctt gtg gca ctg agc tac tgt 624
Leu Gly Ser Val Ile Pro Cys Arg Leu Leu Val Ala Leu Ser Tyr Cys
195 200 205

ttg gcg ttt tgg ggg cac gct gac cac gag ccc tgg gtc agg ttg ttc 672
Leu Ala Phe Trp Gly His Ala Asp His Glu Pro Trp Val Arg Leu Phe
210 215 220

gcc ggc aag att ttt ata ctg tac tta ata att tcc ggc cac att atg 720

Ala Gly Lys Ile Phe Ile Leu Tyr Leu Ile Ile Ser Gly His Ile Met
225 230 235 240

cca cgg aag tca atc tta gaa cag gtg gga acg tcc ggg tac ggt ggc 768
Pro Arg Lys Ser Ile Leu Glu Gln Val Gly Thr Ser Gly Tyr Gly Gly
245 250 255

ttc gtc gag gca gtg tgt cgc gat gtt cgc gcg gta cac ggg att ccg 816
Phe Val Glu Ala Val Cys Arg Asp Val Arg Ala Val His Gly Ile Pro
260 265 270

gcg tgg gat ttt gcg tcc gcg gct ccg gcg tta acg tct caa cag acg 864
Ala Trp Asp Phe Ala Ser Ala Ala Pro Ala Leu Thr Ser Gln Gln Thr
275 280 285

gac tac ttg ttt gcg ttc aat aac agc gta gtg tag 900
Asp Tyr Leu Phe Ala Phe Asn Asn Ser Val Val
290 295 300

<210> 31
<211> 299
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 31
Met Phe Ile Gly Arg Gly Ser Val Tyr Gly Ser Arg Val Ala Thr Ile
1 5 10 15
Glu Gly Ser Lys Tyr Ser Ser Phe Ser Ile Phe Gly Arg Leu Thr Thr
20 25 30
Ser Thr Tyr Pro Pro Thr Tyr Thr Gly Val Met Leu Gly Arg Cys Leu
35 40 45
Arg Glu Pro Lys Glu Met Ser Ala Gly Leu Arg Gly Leu Met Trp Arg
50 55 60
Val Ile Arg Cys Glu Asn Leu Asn Thr Phe Leu Pro Gly Glu Leu Arg
65 70 75 80
Phe Leu His Leu Val Leu Cys Glu Met Tyr Asn Tyr Gly Leu Asn Val
85 90 95
Tyr Leu Leu Lys Glu Ala Ile Ala Asn Thr Gly Thr Arg Asp Asp Ile
100 105 110
Val Leu Gly Arg Lys Val Pro Val Glu Phe Trp Lys Ile Val Tyr Asp
115 120 125
Gly Leu Lys Glu Met Gly Val Ser Asp Ala Thr Leu Leu Ser Glu Thr
130 135 140
Lys Arg Gly Ala Leu Trp Leu Tyr Phe Asn Gly Arg Pro Cys Leu Leu
145 150 155 160
Lys Gly Leu Gly Asp Tyr Val Phe Cys Gln Leu Gly Leu Ser His Ser
165 170 175
Val Arg Val Val Pro Glu Asn Leu Thr Asp Gly Asn Tyr Leu Tyr Asn
180 185 190
Leu Gly Ser Val Ile Pro Cys Arg Leu Leu Val Ala Leu Ser Tyr Cys
195 200 205
Leu Ala Phe Trp Gly His Ala Asp His Glu Pro Trp Val Arg Leu Phe
210 215 220
Ala Gly Lys Ile Phe Ile Leu Tyr Leu Ile Ile Ser Gly His Ile Met
225 230 235 240
Pro Arg Lys Ser Ile Leu Glu Gln Val Gly Thr Ser Gly Tyr Gly Gly
245 250 255
Phe Val Glu Ala Val Cys Arg Asp Val Arg Ala Val His Gly Ile Pro
260 265 270
Ala Trp Asp Phe Ala Ser Ala Ala Pro Ala Leu Thr Ser Gln Gln Thr

```
        275                 280                 285
Asp Tyr Leu Phe Ala Phe Asn Asn Ser Val Val
    290                 295
```

<210> 32
<211> 1644
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1644)

<400> 32

```
atg cgc acg tca gaa aaa tgc tgt atg aga tac cca cga aaa cca gcg    48
Met Arg Thr Ser Glu Lys Cys Cys Met Arg Tyr Pro Arg Lys Pro Ala
1               5                   10                  15

cgg caa att acc gca act ttt tgg gcc ccc cat cca aac aac gtt ctg    96
Arg Gln Ile Thr Ala Thr Phe Trp Ala Pro His Pro Asn Asn Val Leu
            20                  25                  30

ttt att cac aaa cca tct ctg ata gaa gaa aga aga aac gcg ttt gtg   144
Phe Ile His Lys Pro Ser Leu Ile Glu Glu Arg Arg Asn Ala Phe Val
        35                  40                  45

atg cga aat caa cag ctg gcg cta agg gtt cac acg ctg cga aaa aac   192
Met Arg Asn Gln Gln Leu Ala Leu Arg Val His Thr Leu Arg Lys Asn
    50                  55                  60

ctg ctc cgg cta gaa cta gac aac gtg ctg caa acg cac caa cgg gaa   240
Leu Leu Arg Leu Glu Leu Asp Asn Val Leu Gln Thr His Gln Arg Glu
65                  70                  75                  80

aca gaa gtg gtc atg cgc gac cta gaa aca att caa aac atg gtc ggg   288
Thr Glu Val Val Met Arg Asp Leu Glu Thr Ile Gln Asn Met Val Gly
                85                  90                  95

gat ctg cgg tcc ccg ggc agg gaa acc gcc aat gcc caa aca tca cta   336
Asp Leu Arg Ser Pro Gly Arg Glu Thr Ala Asn Ala Gln Thr Ser Leu
            100                 105                 110

aac ccg caa ccc aaa atc gcc ccg cag aca cat gga gac gcg ttc gtt   384
Asn Pro Gln Pro Lys Ile Ala Pro Gln Thr His Gly Asp Ala Phe Val
        115                 120                 125

gtc aca ata gcc ccc gga gac ccg gga ttc acg gta aac cag gat ctc   432
Val Thr Ile Ala Pro Gly Asp Pro Gly Phe Thr Val Asn Gln Asp Leu
    130                 135                 140

cgt ctg gag ctc ctg cca agc ctg tac atg aac caa aac caa tgg ctg   480
Arg Leu Glu Leu Leu Pro Ser Leu Tyr Met Asn Gln Asn Gln Trp Leu
145                 150                 155                 160

cca caa tac ggg ccg tgg tac agt agc ctc aca gat aac gcc atg caa   528
Pro Gln Tyr Gly Pro Trp Tyr Ser Ser Leu Thr Asp Asn Ala Met Gln
                165                 170                 175

cgc agg gtg ttc ccc aga gat ttg cga ggc aca aca aac ttt caa aac   576
```

Arg Arg Val Phe Pro Arg Asp Leu Arg Gly Thr Thr Asn Phe Gln Asn
180 185 190

tcc acg tct cta aag cta atg tct gcc gta atc tca aca gcc gcc tcg 624
Ser Thr Ser Leu Lys Leu Met Ser Ala Val Ile Ser Thr Ala Ala Ser
195 200 205

atc aca caa gac ttt tac gcg gat gtc cgc aac gta tcg gac acg cag 672
Ile Thr Gln Asp Phe Tyr Ala Asp Val Arg Asn Val Ser Asp Thr Gln
210 215 220

gcc gcg ctg tgc ctc ctc aac ggg tac tac tgt cac cgg aca ggt acc 720
Ala Ala Leu Cys Leu Leu Asn Gly Tyr Tyr Cys His Arg Thr Gly Thr
225 230 235 240

ccg ttg ccg cca act cga aac ggc ctc tgg gat aac ctc ggc aca aag 768
Pro Leu Pro Pro Thr Arg Asn Gly Leu Trp Asp Asn Leu Gly Thr Lys
245 250 255

ctg gcg acg ctt gtc tct cac cta aaa caa aac acg aag ggc ctg ggt 816
Leu Ala Thr Leu Val Ser His Leu Lys Gln Asn Thr Lys Gly Leu Gly
260 265 270

ttc gag ttt aca tat tcg aat ccg agg caa cgc gcc tct cta gcg ccc 864
Phe Glu Phe Thr Tyr Ser Asn Pro Arg Gln Arg Ala Ser Leu Ala Pro
275 280 285

cta aac aaa gaa aca aag tac aac gcg gac ttc ttt aca aac cac gtg 912
Leu Asn Lys Glu Thr Lys Tyr Asn Ala Asp Phe Phe Thr Asn His Val
290 295 300

ata tac gcg acg ttg gca cag agc gga ctg cta ccc ggt tcc aag aac 960
Ile Tyr Ala Thr Leu Ala Gln Ser Gly Leu Leu Pro Gly Ser Lys Asn
305 310 315 320

ccc ggt acc ggt cag cca cct gga cca gat ctc gta tac atc ctt gcg 1008
Pro Gly Thr Gly Gln Pro Pro Gly Pro Asp Leu Val Tyr Ile Leu Ala
325 330 335

acc act ttg ttc agc gaa gac gtg cct ccc ttt caa gcg tac caa tgg 1056
Thr Thr Leu Phe Ser Glu Asp Val Pro Pro Phe Gln Ala Tyr Gln Trp
340 345 350

aac ctc cga gcc ggc ctc tcg gca cta ggc tgt tta gtt ttg gtt tac 1104
Asn Leu Arg Ala Gly Leu Ser Ala Leu Gly Cys Leu Val Leu Val Tyr
355 360 365

gtg cta ttg gag ctg gcc caa atc aca ccc aga tcg ccc cac cgg cgc 1152
Val Leu Leu Glu Leu Ala Gln Ile Thr Pro Arg Ser Pro His Arg Arg
370 375 380

ctc aac ctg gcc tcg ctc ctg ggt ggg aga ttt tcc aag gta gaa gac 1200
Leu Asn Leu Ala Ser Leu Leu Gly Gly Arg Phe Ser Lys Val Glu Asp
385 390 395 400

ccg tcc ggg tcc aaa caa tac ctt aaa aag gga caa ctt ttt gat ttc 1248
Pro Ser Gly Ser Lys Gln Tyr Leu Lys Lys Gly Gln Leu Phe Asp Phe
405 410 415

ctg aca gaa aat tac atc tcc cca att ctt tcg cgc gcg cca gac gcc 1296
Leu Thr Glu Asn Tyr Ile Ser Pro Ile Leu Ser Arg Ala Pro Asp Ala

420 425 430

ccc acg tcg ttt ttg ttc ccg ggt gcg tat cta gcg gcg ctg gag gca 1344
Pro Thr Ser Phe Leu Phe Pro Gly Ala Tyr Leu Ala Ala Leu Glu Ala
435 440 445

aag gcc ata tcc cat cta aaa cac act cgt ccg ttc gtc aac ctg acc 1392
Lys Ala Ile Ser His Leu Lys His Thr Arg Pro Phe Val Asn Leu Thr
450 455 460

ggt tcc aga ttt aac gaa atc ttt gac atc ctg aac caa aag ctc acg 1440
Gly Ser Arg Phe Asn Glu Ile Phe Asp Ile Leu Asn Gln Lys Leu Thr
465 470 475 480

ttc agg gac gct ggc agc ctg att cag gct caa acg tcg cta cgt tta 1488
Phe Arg Asp Ala Gly Ser Leu Ile Gln Ala Gln Thr Ser Leu Arg Leu
485 490 495

acg gcc gag gaa ggc ctg gcc gcc atc ctt tcg cac cca tcg ccc ccc 1536
Thr Ala Glu Glu Gly Leu Ala Ala Ile Leu Ser His Pro Ser Pro Pro
500 505 510

ggt ctg gca cac gag ata atg aaa agc caa ttc gga gtt tac gac gac 1584
Gly Leu Ala His Glu Ile Met Lys Ser Gln Phe Gly Val Tyr Asp Asp
515 520 525

tac gac aga gtt tat ttt ttg gtc ctg gga tac ctg ccg gtc gcc acg 1632
Tyr Asp Arg Val Tyr Phe Leu Val Leu Gly Tyr Leu Pro Val Ala Thr
530 535 540

tcg gta gtg tga 1644
Ser Val Val
545

<210> 33
<211> 547
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 33
Met Arg Thr Ser Glu Lys Cys Cys Met Arg Tyr Pro Arg Lys Pro Ala
1 5 10 15
Arg Gln Ile Thr Ala Thr Phe Trp Ala Pro His Pro Asn Asn Val Leu
20 25 30
Phe Ile His Lys Pro Ser Leu Ile Glu Glu Arg Arg Asn Ala Phe Val
35 40 45
Met Arg Asn Gln Gln Leu Ala Leu Arg Val His Thr Leu Arg Lys Asn
50 55 60
Leu Leu Arg Leu Glu Leu Asp Asn Val Leu Gln Thr His Gln Arg Glu
65 70 75 80
Thr Glu Val Val Met Arg Asp Leu Glu Thr Ile Gln Asn Met Val Gly
85 90 95
Asp Leu Arg Ser Pro Gly Arg Glu Thr Ala Asn Ala Gln Thr Ser Leu
100 105 110
Asn Pro Gln Pro Lys Ile Ala Pro Gln Thr His Gly Asp Ala Phe Val
115 120 125
Val Thr Ile Ala Pro Gly Asp Pro Gly Phe Thr Val Asn Gln Asp Leu
130 135 140
Arg Leu Glu Leu Leu Pro Ser Leu Tyr Met Asn Gln Asn Gln Trp Leu
145 150 155 160

```
Pro Gln Tyr Gly Pro Trp Tyr Ser Ser Leu Thr Asp Asn Ala Met Gln
                165                 170                 175
Arg Arg Val Phe Pro Arg Asp Leu Arg Gly Thr Thr Asn Phe Gln Asn
            180                 185                 190
Ser Thr Ser Leu Lys Leu Met Ser Ala Val Ile Ser Thr Ala Ala Ser
        195                 200                 205
Ile Thr Gln Asp Phe Tyr Ala Asp Val Arg Asn Val Ser Asp Thr Gln
    210                 215                 220
Ala Ala Leu Cys Leu Leu Asn Gly Tyr Tyr Cys His Arg Thr Gly Thr
225                 230                 235                 240
Pro Leu Pro Pro Thr Arg Asn Gly Leu Trp Asp Asn Leu Gly Thr Lys
                245                 250                 255
Leu Ala Thr Leu Val Ser His Leu Lys Gln Asn Thr Lys Gly Leu Gly
            260                 265                 270
Phe Glu Phe Thr Tyr Ser Asn Pro Arg Gln Arg Ala Ser Leu Ala Pro
        275                 280                 285
Leu Asn Lys Glu Thr Lys Tyr Asn Ala Asp Phe Phe Thr Asn His Val
    290                 295                 300
Ile Tyr Ala Thr Leu Ala Gln Ser Gly Leu Leu Pro Gly Ser Lys Asn
305                 310                 315                 320
Pro Gly Thr Gly Gln Pro Pro Gly Pro Asp Leu Val Tyr Ile Leu Ala
                325                 330                 335
Thr Thr Leu Phe Ser Glu Asp Val Pro Pro Phe Gln Ala Tyr Gln Trp
            340                 345                 350
Asn Leu Arg Ala Gly Leu Ser Ala Leu Gly Cys Leu Val Leu Val Tyr
        355                 360                 365
Val Leu Leu Glu Leu Ala Gln Ile Thr Pro Arg Ser Pro His Arg Arg
    370                 375                 380
Leu Asn Leu Ala Ser Leu Leu Gly Gly Arg Phe Ser Lys Val Glu Asp
385                 390                 395                 400
Pro Ser Gly Ser Lys Gln Tyr Leu Lys Lys Gly Gln Leu Phe Asp Phe
                405                 410                 415
Leu Thr Glu Asn Tyr Ile Ser Pro Ile Leu Ser Arg Ala Pro Asp Ala
            420                 425                 430
Pro Thr Ser Phe Leu Phe Pro Gly Ala Tyr Leu Ala Ala Leu Glu Ala
        435                 440                 445
Lys Ala Ile Ser His Leu Lys His Thr Arg Pro Phe Val Asn Leu Thr
    450                 455                 460
Gly Ser Arg Phe Asn Glu Ile Phe Asp Ile Leu Asn Gln Lys Leu Thr
465                 470                 475                 480
Phe Arg Asp Ala Gly Ser Leu Ile Gln Ala Gln Thr Ser Leu Arg Leu
                485                 490                 495
Thr Ala Glu Glu Gly Leu Ala Ala Ile Leu Ser His Pro Ser Pro Pro
            500                 505                 510
Gly Leu Ala His Glu Ile Met Lys Ser Gln Phe Gly Val Tyr Asp Asp
        515                 520                 525
Tyr Asp Arg Val Tyr Phe Leu Val Leu Gly Tyr Leu Pro Val Ala Thr
    530                 535                 540
Ser Val Val
545
```

<210> 34
<211> 1053
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1053)

<400> 34
atg gcg ttt gca aat caa tgt aag cac gtc gcg acc ctc gaa gcc ctt 48
Met Ala Phe Ala Asn Gln Cys Lys His Val Ala Thr Leu Glu Ala Leu
1 5 10 15

ccg gcc tca cgc aaa cgg gct gga act cgc gcc cac tta gca gta tac 96
Pro Ala Ser Arg Lys Arg Ala Gly Thr Arg Ala His Leu Ala Val Tyr
20 25 30

cga cgg tta atc aaa cac cgg tcg ctt gat gat att tta aag ttt ctt 144
Arg Arg Leu Ile Lys His Arg Ser Leu Asp Asp Ile Leu Lys Phe Leu
35 40 45

tca att cgc cca aca ctc agg gca acc aaa aac gtt aaa ttt agg att 192
Ser Ile Arg Pro Thr Leu Arg Ala Thr Lys Asn Val Lys Phe Arg Ile
50 55 60

ttt ttc gag gtt tcc cta ggt agg cga ata gca gat tgt gtg ctt acc 240
Phe Phe Glu Val Ser Leu Gly Arg Arg Ile Ala Asp Cys Val Leu Thr
65 70 75 80

gtt aat agc gaa cat caa aaa acg tgc tac gta att gaa ctc aag acg 288
Val Asn Ser Glu His Gln Lys Thr Cys Tyr Val Ile Glu Leu Lys Thr
85 90 95

tgc cta agc gcg gct gta ttt ccc ggc aac gca ata aaa ata tca cag 336
Cys Leu Ser Ala Ala Val Phe Pro Gly Asn Ala Ile Lys Ile Ser Gln
100 105 110

cga tgg cag ggc tta cac caa ctg aca gat tcc gtg gca tat ata ggc 384
Arg Trp Gln Gly Leu His Gln Leu Thr Asp Ser Val Ala Tyr Ile Gly
115 120 125

aga gcc gcg cca aga gga cac gaa aac tgg tcg gta agg ccc tgg ctg 432
Arg Ala Ala Pro Arg Gly His Glu Asn Trp Ser Val Arg Pro Trp Leu
130 135 140

cta ttt aag aac cag aaa aca ctt aaa aca att cat act gag tca tca 480
Leu Phe Lys Asn Gln Lys Thr Leu Lys Thr Ile His Thr Glu Ser Ser
145 150 155 160

gcg ttt ccg cct aca ttt atc aac acc acg tcg gcg gcg ctt aac ggg 528
Ala Phe Pro Pro Thr Phe Ile Asn Thr Thr Ser Ala Ala Leu Asn Gly
165 170 175

ttt ttt agc caa tgg gaa gat gcg cac gtc aga aaa atg ctg tat gag 576
Phe Phe Ser Gln Trp Glu Asp Ala His Val Arg Lys Met Leu Tyr Glu
180 185 190

ata ccc acg aaa acc agc gcg gca aat tac cgc aac ttt ttg ggc ccc 624
Ile Pro Thr Lys Thr Ser Ala Ala Asn Tyr Arg Asn Phe Leu Gly Pro
195 200 205

cca tcc aaa caa cgt tct gtt tat tca caa acc atc tct gat aga aga 672
Pro Ser Lys Gln Arg Ser Val Tyr Ser Gln Thr Ile Ser Asp Arg Arg
210 215 220

aag aag aaa cgc gtt tgt gat gcg aaa tca aca gct ggc gct aag ggt 720
Lys Lys Lys Arg Val Cys Asp Ala Lys Ser Thr Ala Gly Ala Lys Gly
225 230 235 240

tca cac gct gcg aaa aaa cct gct ccg gct aga act aga caa cgt gct 768
Ser His Ala Ala Lys Lys Pro Ala Pro Ala Arg Thr Arg Gln Arg Ala
245 250 255

gca aac gca cca acg gga aac aga agt ggt cat gcg cga cct aga aac 816
Ala Asn Ala Pro Thr Gly Asn Arg Ser Gly His Ala Arg Pro Arg Asn
260 265 270

aat tca aaa cat ggt cgg gga tct gcg gtc ccc ggg cag gga aac cgc 864
Asn Ser Lys His Gly Arg Gly Ser Ala Val Pro Gly Gln Gly Asn Arg
275 280 285

caa tgc cca aac atc act aaa ccc gca acc caa aat cgc ccc gca gac 912
Gln Cys Pro Asn Ile Thr Lys Pro Ala Thr Gln Asn Arg Pro Ala Asp
290 295 300

aca tgg aga cgc gtt cgt tgt cac aat agc ccc cgg aga ccc ggg att 960
Thr Trp Arg Arg Val Arg Cys His Asn Ser Pro Arg Arg Pro Gly Ile
305 310 315 320

cac ggt aaa cca gga tct ccg tct gga gct cct gcc aag cct gta cat 1008
His Gly Lys Pro Gly Ser Pro Ser Gly Ala Pro Ala Lys Pro Val His
325 330 335

gaa cca aaa cca atg gct gcc aca ata cgg gcc gtg gta cag tag 1053
Glu Pro Lys Pro Met Ala Ala Thr Ile Arg Ala Val Val Gln
340 345 350

<210> 35
<211> 350
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 35
Met Ala Phe Ala Asn Gln Cys Lys His Val Ala Thr Leu Glu Ala Leu
1 5 10 15
Pro Ala Ser Arg Lys Arg Ala Gly Thr Arg Ala His Leu Ala Val Tyr
20 25 30
Arg Arg Leu Ile Lys His Arg Ser Leu Asp Asp Ile Leu Lys Phe Leu
35 40 45
Ser Ile Arg Pro Thr Leu Arg Ala Thr Lys Asn Val Lys Phe Arg Ile
50 55 60
Phe Phe Glu Val Ser Leu Gly Arg Arg Ile Ala Asp Cys Val Leu Thr
65 70 75 80
Val Asn Ser Glu His Gln Lys Thr Cys Tyr Val Ile Glu Leu Lys Thr
85 90 95
Cys Leu Ser Ala Ala Val Phe Pro Gly Asn Ala Ile Lys Ile Ser Gln
100 105 110
Arg Trp Gln Gly Leu His Gln Leu Thr Asp Ser Val Ala Tyr Ile Gly
115 120 125
Arg Ala Ala Pro Arg Gly His Glu Asn Trp Ser Val Arg Pro Trp Leu
130 135 140
Leu Phe Lys Asn Gln Lys Thr Leu Lys Thr Ile His Thr Glu Ser Ser
145 150 155 160
Ala Phe Pro Pro Thr Phe Ile Asn Thr Thr Ser Ala Ala Leu Asn Gly
165 170 175
Phe Phe Ser Gln Trp Glu Asp Ala His Val Arg Lys Met Leu Tyr Glu
180 185 190
Ile Pro Thr Lys Thr Ser Ala Ala Asn Tyr Arg Asn Phe Leu Gly Pro

195 200 205
Pro Ser Lys Gln Arg Ser Val Tyr Ser Gln Thr Ile Ser Asp Arg Arg
210 215 220
Lys Lys Lys Arg Val Cys Asp Ala Lys Ser Thr Ala Gly Ala Lys Gly
225 230 235 240
Ser His Ala Ala Lys Lys Pro Ala Pro Ala Arg Thr Arg Gln Arg Ala
245 250 255
Ala Asn Ala Pro Thr Gly Asn Arg Ser Gly His Ala Arg Pro Arg Asn
260 265 270
Asn Ser Lys His Gly Arg Gly Ser Ala Val Pro Gly Gln Gly Asn Arg
275 280 285
Gln Cys Pro Asn Ile Thr Lys Pro Ala Thr Gln Asn Arg Pro Ala Asp
290 295 300
Thr Trp Arg Arg Val Arg Cys His Asn Ser Pro Arg Arg Pro Gly Ile
305 310 315 320
His Gly Lys Pro Gly Ser Pro Ser Gly Ala Pro Ala Lys Pro Val His
325 330 335
Glu Pro Lys Pro Met Ala Ala Thr Ile Arg Ala Val Val Gln
340 345 350

<210> 36
<211> 1674
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1674)

<400> 36
atg gcc gaa gga ggg tct ggc ttt ggc gac gag ctt gtc cgt cag atg 48
Met Ala Glu Gly Gly Ser Gly Phe Gly Asp Glu Leu Val Arg Gln Met
1 5 10 15

cgg gat cgg aag cct aga tgg gac gag tct tcc gac gat act gat gac 96
Arg Asp Arg Lys Pro Arg Trp Asp Glu Ser Ser Asp Asp Thr Asp Asp
20 25 30

gtt gac acc gaa agc acc gat ctt gag tat gat gat gtg ttc ccc gtg 144
Val Asp Thr Glu Ser Thr Asp Leu Glu Tyr Asp Asp Val Phe Pro Val
35 40 45

gta gat acg cac ggc tta atg agc cct gga agt caa aac tat gac gta 192
Val Asp Thr His Gly Leu Met Ser Pro Gly Ser Gln Asn Tyr Asp Val
50 55 60

ccc acg tct ccg tcc gga acg ccg tgg gaa tta ctg cac ccg gac gcc 240
Pro Thr Ser Pro Ser Gly Thr Pro Trp Glu Leu Leu His Pro Asp Ala
65 70 75 80

ctg tat gca cat ccg agg tgc ccg cct aaa agg gcg gtg gtg ccg ggc 288
Leu Tyr Ala His Pro Arg Cys Pro Pro Lys Arg Ala Val Val Pro Gly
85 90 95

ggt ggc gcg cgt ccc aag gtg tcc gcg ttc tcg gct aga ctt caa tat 336
Gly Gly Ala Arg Pro Lys Val Ser Ala Phe Ser Ala Arg Leu Gln Tyr
100 105 110

gtt gga cga cag agt ttt ggg gat aga gag acg cgg cag ctt acc ggg 384

Val Gly Arg Gln Ser Phe Gly Asp Arg Glu Thr Arg Gln Leu Thr Gly
115 120 125

gct cag ttt tct tcc gag agc gag cac gaa tat gcg gaa ata ccg gag 432
Ala Gln Phe Ser Ser Glu Ser Glu His Glu Tyr Ala Glu Ile Pro Glu
130 135 140

cgt act act aca cgc ccg gtt gaa agc ggt gac aag aga aat ttc acc 480
Arg Thr Thr Thr Arg Pro Val Glu Ser Gly Asp Lys Arg Asn Phe Thr
145 150 155 160

tcc ggg cgt agg ggt gca atc tcg gga ccc tcg tcg aca aaa cct agt 528
Ser Gly Arg Arg Gly Ala Ile Ser Gly Pro Ser Ser Thr Lys Pro Ser
165 170 175

cac ggt gcg ggg tta aca cgt aag act aaa acg tcg ctg agt gtt agt 576
His Gly Ala Gly Leu Thr Arg Lys Thr Lys Thr Ser Leu Ser Val Ser
180 185 190

ctt aaa aac ttg ctg cgg ata aag gac gat gat gct aag gtt gat gta 624
Leu Lys Asn Leu Leu Arg Ile Lys Asp Asp Asp Ala Lys Val Asp Val
195 200 205

ccc agg ccg gtt acc gtg ccg gtt cat ctc atg cag ccg cat cct atg 672
Pro Arg Pro Val Thr Val Pro Val His Leu Met Gln Pro His Pro Met
210 215 220

acg gaa tat aga aat gcc ttt tta atc tat cta gag ggg gtc atg ggt 720
Thr Glu Tyr Arg Asn Ala Phe Leu Ile Tyr Leu Glu Gly Val Met Gly
225 230 235 240

gtt gga aaa aca acg cta ttg aat tca atg act ggc atg gtg ccg cag 768
Val Gly Lys Thr Thr Leu Leu Asn Ser Met Thr Gly Met Val Pro Gln
245 250 255

gaa aac gtc tta agc tgt ccc gag ccc atg aaa ttt tgg acg tgt gtt 816
Glu Asn Val Leu Ser Cys Pro Glu Pro Met Lys Phe Trp Thr Cys Val
260 265 270

tat tca aat tgc ctt aaa gaa cag cgc agc ata gtt aag caa ggc acc 864
Tyr Ser Asn Cys Leu Lys Glu Gln Arg Ser Ile Val Lys Gln Gly Thr
275 280 285

cac ggg aaa ttg atc act tct gct cgc gta tac gcg tgt cag agc aag 912
His Gly Lys Leu Ile Thr Ser Ala Arg Val Tyr Ala Cys Gln Ser Lys
290 295 300

ttt gcg cta ccg ttt cgt gcg aca gcc gcc ggc atc ggt cgc aac ctg 960
Phe Ala Leu Pro Phe Arg Ala Thr Ala Ala Gly Ile Gly Arg Asn Leu
305 310 315 320

caa ccg tgg ctg gtg gga aac ggt agc aca aag cca gcg aat tgg att 1008
Gln Pro Trp Leu Val Gly Asn Gly Ser Thr Lys Pro Ala Asn Trp Ile
325 330 335

gtt ttt gac aga cac cta ctg tcc gca acg gtg gtt ttt ccg ctg gtt 1056
Val Phe Asp Arg His Leu Leu Ser Ala Thr Val Val Phe Pro Leu Val
340 345 350

cac gta aag tac aac agg cta acg ccg gat cat ctg ttt caa atc cta 1104
His Val Lys Tyr Asn Arg Leu Thr Pro Asp His Leu Phe Gln Ile Leu

355 360 365

tct ctt ttc tcg gca cac gac ggc gat gtg gtc gtt ttg cta acg ctc 1152
Ser Leu Phe Ser Ala His Asp Gly Asp Val Val Val Leu Leu Thr Leu
370 375 380

aac agt tcg gag gcg cac agg cgc att caa agt cga ggc cgt aag gaa 1200
Asn Ser Ser Glu Ala His Arg Arg Ile Gln Ser Arg Gly Arg Lys Glu
385 390 395 400

gaa aaa gga atc acg caa aac tac ttg cga cag gta gcg tgg gcg tac 1248
Glu Lys Gly Ile Thr Gln Asn Tyr Leu Arg Gln Val Ala Trp Ala Tyr
405 410 415

cat gcc gtg ttc tgt acg tgg gtg atg atg caa tat ctc aca ccg gag 1296
His Ala Val Phe Cys Thr Trp Val Met Met Gln Tyr Leu Thr Pro Glu
420 425 430

caa atg gtt cag ctg tgt gta caa act gtg tcc ata gag gac ata tgc 1344
Gln Met Val Gln Leu Cys Val Gln Thr Val Ser Ile Glu Asp Ile Cys
435 440 445

aat atg aat tcc aga ttg act cat cgg ttt cta acc cta aca aaa cta 1392
Asn Met Asn Ser Arg Leu Thr His Arg Phe Leu Thr Leu Thr Lys Leu
450 455 460

cat gaa cag agt atg ata ccc atg gtg gca gaa atg tta gta tcg gtt 1440
His Glu Gln Ser Met Ile Pro Met Val Ala Glu Met Leu Val Ser Val
465 470 475 480

aaa gaa cac gtg acc tta atg gag gtc tgt ttg gga ctc ttt aaa gag 1488
Lys Glu His Val Thr Leu Met Glu Val Cys Leu Gly Leu Phe Lys Glu
485 490 495

cta cga aag ctt caa att tta att gtt gac gca gga gaa cat cta gat 1536
Leu Arg Lys Leu Gln Ile Leu Ile Val Asp Ala Gly Glu His Leu Asp
500 505 510

gat gcg tgt ggc ctc tgg gga aat att tat ggg cag gtg atg tca aat 1584
Asp Ala Cys Gly Leu Trp Gly Asn Ile Tyr Gly Gln Val Met Ser Asn
515 520 525

gag gct att aaa cca cga gca gtg aac tgg cca gct ctt gaa agc tac 1632
Glu Ala Ile Lys Pro Arg Ala Val Asn Trp Pro Ala Leu Glu Ser Tyr
530 535 540

att caa acg cta acc aaa ttg gaa ggc aat ggc gcg tat tag 1674
Ile Gln Thr Leu Thr Lys Leu Glu Gly Asn Gly Ala Tyr
545 550 555

<210> 37
<211> 557
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 37
Met Ala Glu Gly Gly Ser Gly Phe Gly Asp Glu Leu Val Arg Gln Met
1 5 10 15
Arg Asp Arg Lys Pro Arg Trp Asp Glu Ser Ser Asp Asp Thr Asp Asp
20 25 30

Val Asp Thr Glu Ser Thr Asp Leu Glu Tyr Asp Asp Val Phe Pro Val
35 40 45

Val Asp Thr His Gly Leu Met Ser Pro Gly Ser Gln Asn Tyr Asp Val
50 55 60

Pro Thr Ser Pro Ser Gly Thr Pro Trp Glu Leu Leu His Pro Asp Ala
65 70 75 80

Leu Tyr Ala His Pro Arg Cys Pro Pro Lys Arg Ala Val Val Pro Gly
85 90 95

Gly Gly Ala Arg Pro Lys Val Ser Ala Phe Ser Ala Arg Leu Gln Tyr
100 105 110

Val Gly Arg Gln Ser Phe Gly Asp Arg Glu Thr Arg Gln Leu Thr Gly
115 120 125

Ala Gln Phe Ser Ser Glu Ser Glu His Glu Tyr Ala Glu Ile Pro Glu
130 135 140

Arg Thr Thr Thr Arg Pro Val Glu Ser Gly Asp Lys Arg Asn Phe Thr
145 150 155 160

Ser Gly Arg Arg Gly Ala Ile Ser Gly Pro Ser Ser Thr Lys Pro Ser
165 170 175

His Gly Ala Gly Leu Thr Arg Lys Thr Lys Thr Ser Leu Ser Val Ser
180 185 190

Leu Lys Asn Leu Leu Arg Ile Lys Asp Asp Asp Ala Lys Val Asp Val
195 200 205

Pro Arg Pro Val Thr Val Pro Val His Leu Met Gln Pro His Pro Met
210 215 220

Thr Glu Tyr Arg Asn Ala Phe Leu Ile Tyr Leu Glu Gly Val Met Gly
225 230 235 240

Val Gly Lys Thr Thr Leu Leu Asn Ser Met Thr Gly Met Val Pro Gln
245 250 255

Glu Asn Val Leu Ser Cys Pro Glu Pro Met Lys Phe Trp Thr Cys Val
260 265 270

Tyr Ser Asn Cys Leu Lys Glu Gln Arg Ser Ile Val Lys Gln Gly Thr
275 280 285

His Gly Lys Leu Ile Thr Ser Ala Arg Val Tyr Ala Cys Gln Ser Lys
290 295 300

Phe Ala Leu Pro Phe Arg Ala Thr Ala Ala Gly Ile Gly Arg Asn Leu
305 310 315 320

Gln Pro Trp Leu Val Gly Asn Gly Ser Thr Lys Pro Ala Asn Trp Ile
325 330 335

Val Phe Asp Arg His Leu Leu Ser Ala Thr Val Val Phe Pro Leu Val
340 345 350

His Val Lys Tyr Asn Arg Leu Thr Pro Asp His Leu Phe Gln Ile Leu
355 360 365

Ser Leu Phe Ser Ala His Asp Gly Asp Val Val Val Leu Leu Thr Leu
370 375 380

Asn Ser Ser Glu Ala His Arg Arg Ile Gln Ser Arg Gly Arg Lys Glu
385 390 395 400

Glu Lys Gly Ile Thr Gln Asn Tyr Leu Arg Gln Val Ala Trp Ala Tyr
405 410 415

His Ala Val Phe Cys Thr Trp Val Met Met Gln Tyr Leu Thr Pro Glu
420 425 430

Gln Met Val Gln Leu Cys Val Gln Thr Val Ser Ile Glu Asp Ile Cys
435 440 445

Asn Met Asn Ser Arg Leu Thr His Arg Phe Leu Thr Leu Thr Lys Leu
450 455 460

His Glu Gln Ser Met Ile Pro Met Val Ala Glu Met Leu Val Ser Val
465 470 475 480

Lys Glu His Val Thr Leu Met Glu Val Cys Leu Gly Leu Phe Lys Glu
485 490 495

Leu Arg Lys Leu Gln Ile Leu Ile Val Asp Ala Gly Glu His Leu Asp
500 505 510

Asp Ala Cys Gly Leu Trp Gly Asn Ile Tyr Gly Gln Val Met Ser Asn

515 520 525
Glu Ala Ile Lys Pro Arg Ala Val Asn Trp Pro Ala Leu Glu Ser Tyr
530 535 540
Ile Gln Thr Leu Thr Lys Leu Glu Gly Asn Gly Ala Tyr
545 550 555

<210> 38
<211> 2115
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2115)

<400> 38
atg gcg cgt att agt ttt ata ttt ttt ttt act ata ata agg tgt tct 48
Met Ala Arg Ile Ser Phe Ile Phe Phe Phe Thr Ile Ile Arg Cys Ser
1 5 10 15

gta acg gac aaa tac gtg tat gat gaa aaa tca aat gta gag ttg gaa 96
Val Thr Asp Lys Tyr Val Tyr Asp Glu Lys Ser Asn Val Glu Leu Glu
20 25 30

ttt aat gga aca att tac caa att aat tgg aga aac gta agc aaa gaa 144
Phe Asn Gly Thr Ile Tyr Gln Ile Asn Trp Arg Asn Val Ser Lys Glu
35 40 45

tta aca tct att gta atg gaa gat gcg tgg tac gat tct tta ctg ctg 192
Leu Thr Ser Ile Val Met Glu Asp Ala Trp Tyr Asp Ser Leu Leu Leu
50 55 60

gaa cca ttg tct gtg act ttg gaa aaa aga aaa agt ttg ctg cga tca 240
Glu Pro Leu Ser Val Thr Leu Glu Lys Arg Lys Ser Leu Leu Arg Ser
65 70 75 80

agt ata gtt aat gtg cat aat aat gac tat aca ttt tgt aaa tct tca 288
Ser Ile Val Asn Val His Asn Asn Asp Tyr Thr Phe Cys Lys Ser Ser
85 90 95

agt gac cat gta att aac ctt acg gtg gat ttt aat tac agt tct ttg 336
Ser Asp His Val Ile Asn Leu Thr Val Asp Phe Asn Tyr Ser Ser Leu
100 105 110

cca ggt ttt act ggg aat ttt aat gta atg aca cat gct tta aca caa 384
Pro Gly Phe Thr Gly Asn Phe Asn Val Met Thr His Ala Leu Thr Gln
115 120 125

ggc gtt tta tta acg aaa cgt gaa tta ttt acc aac tca aca aac att 432
Gly Val Leu Leu Thr Lys Arg Glu Leu Phe Thr Asn Ser Thr Asn Ile
130 135 140

atg gat ctt ttt tac gcg gaa aaa ata aat gca gaa atg ttc aag att 480
Met Asp Leu Phe Tyr Ala Glu Lys Ile Asn Ala Glu Met Phe Lys Ile
145 150 155 160

act ttt gat tat tcc aat gta ata att tcc ggc att ata aca gaa aat 528
Thr Phe Asp Tyr Ser Asn Val Ile Ile Ser Gly Ile Ile Thr Glu Asn
165 170 175

tgg ata tta gta tca gtt aca aat tca agt gta aaa tca aat atg caa 576
Trp Ile Leu Val Ser Val Thr Asn Ser Ser Val Lys Ser Asn Met Gln
180 185 190

tgt gtt gca ctg ttg ttt ggt gtt cct agc aca ttt cca gcg tta aaa 624
Cys Val Ala Leu Leu Phe Gly Val Pro Ser Thr Phe Pro Ala Leu Lys
195 200 205

ggc tat gta tct tat cgt gat tta ttg gtc gtt aaa aat tca aac tat 672
Gly Tyr Val Ser Tyr Arg Asp Leu Leu Val Val Lys Asn Ser Asn Tyr
210 215 220

gct ctg ggt gta att gcg ccg aaa tct tat aat act ctt gat ttg gca 720
Ala Leu Gly Val Ile Ala Pro Lys Ser Tyr Asn Thr Leu Asp Leu Ala
225 230 235 240

ttt tta cca aaa aat ttt aca gaa atg ttt gtt tca gtt ata gat agt 768
Phe Leu Pro Lys Asn Phe Thr Glu Met Phe Val Ser Val Ile Asp Ser
245 250 255

ccg ctt aat gcc att gat tat ctc aaa gga aaa ctt ttg gct att gaa 816
Pro Leu Asn Ala Ile Asp Tyr Leu Lys Gly Lys Leu Leu Ala Ile Glu
260 265 270

gca aag ggt gcc tgt caa aat cct tca aat gaa aat gat atc tta tct 864
Ala Lys Gly Ala Cys Gln Asn Pro Ser Asn Glu Asn Asp Ile Leu Ser
275 280 285

ttc ttt ttt gaa gtt act gca gta aat ttt tta ttt ata aaa aat ctt 912
Phe Phe Phe Glu Val Thr Ala Val Asn Phe Leu Phe Ile Lys Asn Leu
290 295 300

caa aag cag caa tta gtt aac gtt ggg tgt gtt gta cgg cat gtt gca 960
Gln Lys Gln Gln Leu Val Asn Val Gly Cys Val Val Arg His Val Ala
305 310 315 320

gcg tta gaa agt tta atg cat tta tta agg ctc tgc tac cca acg ttt 1008
Ala Leu Glu Ser Leu Met His Leu Leu Arg Leu Cys Tyr Pro Thr Phe
325 330 335

aag tta tat gag cta aat ttg gaa acg ttg tcg cat att gca gag tcc 1056
Lys Leu Tyr Glu Leu Asn Leu Glu Thr Leu Ser His Ile Ala Glu Ser
340 345 350

cag gtt ttt aac ttg ccg gca aat agc atg tta tcg tta tca gta aat 1104
Gln Val Phe Asn Leu Pro Ala Asn Ser Met Leu Ser Leu Ser Val Asn
355 360 365

gat caa gag gtc gta ttt tca atg ttt aaa att gta tat aat aca cct 1152
Asp Gln Glu Val Val Phe Ser Met Phe Lys Ile Val Tyr Asn Thr Pro
370 375 380

aag gtt ggt ggt aaa att cta aac gaa att gta tat att act aat tac 1200
Lys Val Gly Gly Lys Ile Leu Asn Glu Ile Val Tyr Ile Thr Asn Tyr
385 390 395 400

atg tac acc aaa tac agt gaa aat tat cag cta acc aac aca ttt agg 1248
Met Tyr Thr Lys Tyr Ser Glu Asn Tyr Gln Leu Thr Asn Thr Phe Arg
405 410 415

cgg aat gta atg aac atg tat gaa gtt tta aca aca att aaa tta aac 1296
Arg Asn Val Met Asn Met Tyr Glu Val Leu Thr Thr Ile Lys Leu Asn
420 425 430

gtc aca gat tct tct gtg ttt tat ccg tat ata ttg ttc aca tcc atg 1344
Val Thr Asp Ser Ser Val Phe Tyr Pro Tyr Ile Leu Phe Thr Ser Met
435 440 445

tgt aat aat gta gag ata tca tat atg att aat caa ata gca aag ccc 1392
Cys Asn Asn Val Glu Ile Ser Tyr Met Ile Asn Gln Ile Ala Lys Pro
450 455 460

gat gac atc acc ata ttt cgc gtg ttt tcg cct tgt ttt tta agc tta 1440
Asp Asp Ile Thr Ile Phe Arg Val Phe Ser Pro Cys Phe Leu Ser Leu
465 470 475 480

agg ttt gat tta gat gaa aac aag cta cga tcc gac gca ccg cag acg 1488
Arg Phe Asp Leu Asp Glu Asn Lys Leu Arg Ser Asp Ala Pro Gln Thr
485 490 495

tca aaa aga acc ggc tcg gaa ctg gca caa gga gcg tct ggt ttt tgg 1536
Ser Lys Arg Thr Gly Ser Glu Leu Ala Gln Gly Ala Ser Gly Phe Trp
500 505 510

cgg ctg ttg cac gct ttt cac gcc acg cgt ata aac gag ttt tca gtc 1584
Arg Leu Leu His Ala Phe His Ala Thr Arg Ile Asn Glu Phe Ser Val
515 520 525

ata aat tgt acc agg ctg gca tgg aag caa gta acc gcg ctg atg ccg 1632
Ile Asn Cys Thr Arg Leu Ala Trp Lys Gln Val Thr Ala Leu Met Pro
530 535 540

ctt acg aac atc acg tac gtc ata agc tcc gtg cgc cct gat cac gct 1680
Leu Thr Asn Ile Thr Tyr Val Ile Ser Ser Val Arg Pro Asp His Ala
545 550 555 560

cgc gtt tac gag gtt tcg gag gtg ttc ctt aac agc gcg atg ttt gtc 1728
Arg Val Tyr Glu Val Ser Glu Val Phe Leu Asn Ser Ala Met Phe Val
565 570 575

tcg gcc gtt tat cca aat tgt tct cat ttt aca ccg ccc ggt acc gcc 1776
Ser Ala Val Tyr Pro Asn Cys Ser His Phe Thr Pro Pro Gly Thr Ala
580 585 590

cta cac att ccg ata ttg tat aac ttt tcc gcc ccg agg ata gga tgt 1824
Leu His Ile Pro Ile Leu Tyr Asn Phe Ser Ala Pro Arg Ile Gly Cys
595 600 605

ccg ctg tgt gat tcg att gtt ttg agc tat gat gag aac cag gga cta 1872
Pro Leu Cys Asp Ser Ile Val Leu Ser Tyr Asp Glu Asn Gln Gly Leu
610 615 620

cag acc atg atg tac gtg tca aac ccc aca gtt caa gcc aac ctg ttt 1920
Gln Thr Met Met Tyr Val Ser Asn Pro Thr Val Gln Ala Asn Leu Phe
625 630 635 640

tcc ccg tat tct cct ttc ttt gac aac gat aat ttt cat atc cat tat 1968
Ser Pro Tyr Ser Pro Phe Phe Asp Asn Asp Asn Phe His Ile His Tyr
645 650 655

ttg tgg tta atg aac aac ggt acc gtt gtg gaa att cgg gga ttg tac 2016

Leu Trp Leu Met Asn Asn Gly Thr Val Val Glu Ile Arg Gly Leu Tyr
660 665 670

aga aga cac gca ctc agt gct att gca tta gta ttt gct ttt att gga 2064
Arg Arg His Ala Leu Ser Ala Ile Ala Leu Val Phe Ala Phe Ile Gly
675 680 685

aca atg tct gct ctg tat ttt ctg ttt aag ctg ttt tcc atc ctg gcc 2112
Thr Met Ser Ala Leu Tyr Phe Leu Phe Lys Leu Phe Ser Ile Leu Ala
690 695 700

taa 2115

705

<210> 39
<211> 704
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 39
Met Ala Arg Ile Ser Phe Ile Phe Phe Phe Thr Ile Ile Arg Cys Ser
1 5 10 15
Val Thr Asp Lys Tyr Val Tyr Asp Glu Lys Ser Asn Val Glu Leu Glu
20 25 30
Phe Asn Gly Thr Ile Tyr Gln Ile Asn Trp Arg Asn Val Ser Lys Glu
35 40 45
Leu Thr Ser Ile Val Met Glu Asp Ala Trp Tyr Asp Ser Leu Leu Leu
50 55 60
Glu Pro Leu Ser Val Thr Leu Glu Lys Arg Lys Ser Leu Leu Arg Ser
65 70 75 80
Ser Ile Val Asn Val His Asn Asn Asp Tyr Thr Phe Cys Lys Ser Ser
85 90 95
Ser Asp His Val Ile Asn Leu Thr Val Asp Phe Asn Tyr Ser Ser Leu
100 105 110
Pro Gly Phe Thr Gly Asn Phe Asn Val Met Thr His Ala Leu Thr Gln
115 120 125
Gly Val Leu Leu Thr Lys Arg Glu Leu Phe Thr Asn Ser Thr Asn Ile
130 135 140
Met Asp Leu Phe Tyr Ala Glu Lys Ile Asn Ala Glu Met Phe Lys Ile
145 150 155 160
Thr Phe Asp Tyr Ser Asn Val Ile Ile Ser Gly Ile Ile Thr Glu Asn
165 170 175
Trp Ile Leu Val Ser Val Thr Asn Ser Ser Val Lys Ser Asn Met Gln
180 185 190
Cys Val Ala Leu Leu Phe Gly Val Pro Ser Thr Phe Pro Ala Leu Lys
195 200 205
Gly Tyr Val Ser Tyr Arg Asp Leu Leu Val Val Lys Asn Ser Asn Tyr
210 215 220
Ala Leu Gly Val Ile Ala Pro Lys Ser Tyr Asn Thr Leu Asp Leu Ala
225 230 235 240
Phe Leu Pro Lys Asn Phe Thr Glu Met Phe Val Ser Val Ile Asp Ser
245 250 255
Pro Leu Asn Ala Ile Asp Tyr Leu Lys Gly Lys Leu Leu Ala Ile Glu
260 265 270
Ala Lys Gly Ala Cys Gln Asn Pro Ser Asn Glu Asn Asp Ile Leu Ser
275 280 285
Phe Phe Phe Glu Val Thr Ala Val Asn Phe Leu Phe Ile Lys Asn Leu
290 295 300
Gln Lys Gln Gln Leu Val Asn Val Gly Cys Val Val Arg His Val Ala

305 310 315 320
Ala Leu Glu Ser Leu Met His Leu Leu Arg Leu Cys Tyr Pro Thr Phe
325 330 335
Lys Leu Tyr Glu Leu Asn Leu Glu Thr Leu Ser His Ile Ala Glu Ser
340 345 350
Gln Val Phe Asn Leu Pro Ala Asn Ser Met Leu Ser Leu Ser Val Asn
355 360 365
Asp Gln Glu Val Val Phe Ser Met Phe Lys Ile Val Tyr Asn Thr Pro
370 375 380
Lys Val Gly Gly Lys Ile Leu Asn Glu Ile Val Tyr Ile Thr Asn Tyr
385 390 395 400
Met Tyr Thr Lys Tyr Ser Glu Asn Tyr Gln Leu Thr Asn Thr Phe Arg
405 410 415
Arg Asn Val Met Asn Met Tyr Glu Val Leu Thr Thr Ile Lys Leu Asn
420 425 430
Val Thr Asp Ser Ser Val Phe Tyr Pro Tyr Ile Leu Phe Thr Ser Met
435 440 445
Cys Asn Asn Val Glu Ile Ser Tyr Met Ile Asn Gln Ile Ala Lys Pro
450 455 460
Asp Asp Ile Thr Ile Phe Arg Val Phe Ser Pro Cys Phe Leu Ser Leu
465 470 475 480
Arg Phe Asp Leu Asp Glu Asn Lys Leu Arg Ser Asp Ala Pro Gln Thr
485 490 495
Ser Lys Arg Thr Gly Ser Glu Leu Ala Gln Gly Ala Ser Gly Phe Trp
500 505 510
Arg Leu Leu His Ala Phe His Ala Thr Arg Ile Asn Glu Phe Ser Val
515 520 525
Ile Asn Cys Thr Arg Leu Ala Trp Lys Gln Val Thr Ala Leu Met Pro
530 535 540
Leu Thr Asn Ile Thr Tyr Val Ile Ser Ser Val Arg Pro Asp His Ala
545 550 555 560
Arg Val Tyr Glu Val Ser Glu Val Phe Leu Asn Ser Ala Met Phe Val
565 570 575
Ser Ala Val Tyr Pro Asn Cys Ser His Phe Thr Pro Pro Gly Thr Ala
580 585 590
Leu His Ile Pro Ile Leu Tyr Asn Phe Ser Ala Pro Arg Ile Gly Cys
595 600 605
Pro Leu Cys Asp Ser Ile Val Leu Ser Tyr Asp Glu Asn Gln Gly Leu
610 615 620
Gln Thr Met Met Tyr Val Ser Asn Pro Thr Val Gln Ala Asn Leu Phe
625 630 635 640
Ser Pro Tyr Ser Pro Phe Phe Asp Asn Asp Asn Phe His Ile His Tyr
645 650 655
Leu Trp Leu Met Asn Asn Gly Thr Val Val Glu Ile Arg Gly Leu Tyr
660 665 670
Arg Arg His Ala Leu Ser Ala Ile Ala Leu Val Phe Ala Phe Ile Gly
675 680 685
Thr Met Ser Ala Leu Tyr Phe Leu Phe Lys Leu Phe Ser Ile Leu Ala
690 695 700

<210> 40
<211> 1209
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1209)

<400> 40
atg ata aaa atc tca gat tta aag gct cgc tta gtt ggg ggg gcg gtg 48
Met Ile Lys Ile Ser Asp Leu Lys Ala Arg Leu Val Gly Gly Ala Val
1 5 10 15

cag cta tcc aac ggc gag tac gtc tgt cat gtc gtg tac tcg tcg gca 96
Gln Leu Ser Asn Gly Glu Tyr Val Cys His Val Val Tyr Ser Ser Ala
20 25 30

ctt gcc gca atg gtc gga ctt ccg gga ccg gcc gta ccc ctt ccc ctg 144
Leu Ala Ala Met Val Gly Leu Pro Gly Pro Ala Val Pro Leu Pro Leu
35 40 45

ttg ttt aaa aaa ttc agc acc atc tat tca aac atg atg cca ttg tac 192
Leu Phe Lys Lys Phe Ser Thr Ile Tyr Ser Asn Met Met Pro Leu Tyr
50 55 60

gcc cca aaa cgt ccg gaa tta tca atg ctg cga att atg gtg tca ccc 240
Ala Pro Lys Arg Pro Glu Leu Ser Met Leu Arg Ile Met Val Ser Pro
65 70 75 80

cat ccc tac gcg cta aac agc tgc ttg tgc gtg gga acg gac gaa ggc 288
His Pro Tyr Ala Leu Asn Ser Cys Leu Cys Val Gly Thr Asp Glu Gly
85 90 95

gag aga ggc gta tca ttg ttc aga gac cca gta att aga agt tct gac 336
Glu Arg Gly Val Ser Leu Phe Arg Asp Pro Val Ile Arg Ser Ser Asp
100 105 110

ttt gag gac aca cca att acc gta aac tca aaa tta gtg att gcg agt 384
Phe Glu Asp Thr Pro Ile Thr Val Asn Ser Lys Leu Val Ile Ala Ser
115 120 125

aac tct ctt ttt ttg cac tgc agg cca ttt tcc gtc ccg gcg acc gtc 432
Asn Ser Leu Phe Leu His Cys Arg Pro Phe Ser Val Pro Ala Thr Val
130 135 140

aag acc cca ccg gta act ctt aca aat aac aaa cag ata acc ata aac 480
Lys Thr Pro Pro Val Thr Leu Thr Asn Asn Lys Gln Ile Thr Ile Asn
145 150 155 160

gag ctg gca aac acc acg caa gaa tac gat cca aat gcg cca ccg acc 528
Glu Leu Ala Asn Thr Thr Gln Glu Tyr Asp Pro Asn Ala Pro Pro Thr
165 170 175

tta tgt agc gcg ctt ccg ccg gat aat aaa aaa ctt cga tcc atc ctc 576
Leu Cys Ser Ala Leu Pro Pro Asp Asn Lys Lys Leu Arg Ser Ile Leu
180 185 190

aag caa ccg cca gcg acc agc gag tca aat gta cag tcg gac tgt tta 624
Lys Gln Pro Pro Ala Thr Ser Glu Ser Asn Val Gln Ser Asp Cys Leu
195 200 205

ctg gca gac ata ttt ttt gcc atg ggt tcg cgc cag cca caa ata ggc 672
Leu Ala Asp Ile Phe Phe Ala Met Gly Ser Arg Gln Pro Gln Ile Gly
210 215 220

gag agc ccg att acc gcg ttt aac acc gtg aca att atg cag cgc gcc 720
Glu Ser Pro Ile Thr Ala Phe Asn Thr Val Thr Ile Met Gln Arg Ala
225 230 235 240

aac aac tct ata atg ttt tta cct aat tta aaa cta aag cct atc caa 768
Asn Asn Ser Ile Met Phe Leu Pro Asn Leu Lys Leu Lys Pro Ile Gln
245 250 255

cac ctt ttt tta aag cac gtc cta ttg cag cgt ttg ggt ttg gaa aat 816
His Leu Phe Leu Lys His Val Leu Leu Gln Arg Leu Gly Leu Glu Asn
260 265 270

att ctg ttt cac ttt aaa atg cta tac gcg aat acg tgt aag gcg gcg 864
Ile Leu Phe His Phe Lys Met Leu Tyr Ala Asn Thr Cys Lys Ala Ala
275 280 285

ggc ccg tat cag aga gaa tac ttc gaa tcc atg ctc tcg cgc gta aaa 912
Gly Pro Tyr Gln Arg Glu Tyr Phe Glu Ser Met Leu Ser Arg Val Lys
290 295 300

cag cgc ctg gaa gac atg gtg ttt tgt tta aac tcc ata gag tcc cac 960
Gln Arg Leu Glu Asp Met Val Phe Cys Leu Asn Ser Ile Glu Ser His
305 310 315 320

gac ttt cag aag gat ttt cga gtt ttg tcg cgg gct ccg cag cga tta 1008
Asp Phe Gln Lys Asp Phe Arg Val Leu Ser Arg Ala Pro Gln Arg Leu
325 330 335

tta acc gcg acc gac aaa tat ttt tta atg ttt cct ccc caa aat cga 1056
Leu Thr Ala Thr Asp Lys Tyr Phe Leu Met Phe Pro Pro Gln Asn Arg
340 345 350

gag ctg gcc att caa gtt ggc gca gag gtg att gaa tca ata tgt gac 1104
Glu Leu Ala Ile Gln Val Gly Ala Glu Val Ile Glu Ser Ile Cys Asp
355 360 365

ggc acg cct cta tcc gag gtt ctc gct aat ctc agt ccc cgc gta acc 1152
Gly Thr Pro Leu Ser Glu Val Leu Ala Asn Leu Ser Pro Arg Val Thr
370 375 380

atc caa aaa gaa acg gga aac aac ctt tta aag ttt tac gct tta tta 1200
Ile Gln Lys Glu Thr Gly Asn Asn Leu Leu Lys Phe Tyr Ala Leu Leu
385 390 395 400

aca gtt tag 1209
Thr Val

<210> 41
<211> 402
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 41
Met Ile Lys Ile Ser Asp Leu Lys Ala Arg Leu Val Gly Gly Ala Val
1 5 10 15
Gln Leu Ser Asn Gly Glu Tyr Val Cys His Val Val Tyr Ser Ser Ala
20 25 30
Leu Ala Ala Met Val Gly Leu Pro Gly Pro Ala Val Pro Leu Pro Leu
35 40 45
Leu Phe Lys Lys Phe Ser Thr Ile Tyr Ser Asn Met Met Pro Leu Tyr
50 55 60
Ala Pro Lys Arg Pro Glu Leu Ser Met Leu Arg Ile Met Val Ser Pro
65 70 75 80
His Pro Tyr Ala Leu Asn Ser Cys Leu Cys Val Gly Thr Asp Glu Gly

85 90 95
Glu Arg Gly Val Ser Leu Phe Arg Asp Pro Val Ile Arg Ser Ser Asp
100 105 110
Phe Glu Asp Thr Pro Ile Thr Val Asn Ser Lys Leu Val Ile Ala Ser
115 120 125
Asn Ser Leu Phe Leu His Cys Arg Pro Phe Ser Val Pro Ala Thr Val
130 135 140
Lys Thr Pro Pro Val Thr Leu Thr Asn Asn Lys Gln Ile Thr Ile Asn
145 150 155 160
Glu Leu Ala Asn Thr Thr Gln Glu Tyr Asp Pro Asn Ala Pro Pro Thr
165 170 175
Leu Cys Ser Ala Leu Pro Pro Asp Asn Lys Lys Leu Arg Ser Ile Leu
180 185 190
Lys Gln Pro Pro Ala Thr Ser Glu Ser Asn Val Gln Ser Asp Cys Leu
195 200 205
Leu Ala Asp Ile Phe Phe Ala Met Gly Ser Arg Gln Pro Gln Ile Gly
210 215 220
Glu Ser Pro Ile Thr Ala Phe Asn Thr Val Thr Ile Met Gln Arg Ala
225 230 235 240
Asn Asn Ser Ile Met Phe Leu Pro Asn Leu Lys Leu Lys Pro Ile Gln
245 250 255
His Leu Phe Leu Lys His Val Leu Leu Gln Arg Leu Gly Leu Glu Asn
260 265 270
Ile Leu Phe His Phe Lys Met Leu Tyr Ala Asn Thr Cys Lys Ala Ala
275 280 285
Gly Pro Tyr Gln Arg Glu Tyr Phe Glu Ser Met Leu Ser Arg Val Lys
290 295 300
Gln Arg Leu Glu Asp Met Val Phe Cys Leu Asn Ser Ile Glu Ser His
305 310 315 320
Asp Phe Gln Lys Asp Phe Arg Val Leu Ser Arg Ala Pro Gln Arg Leu
325 330 335
Leu Thr Ala Thr Asp Lys Tyr Phe Leu Met Phe Pro Pro Gln Asn Arg
340 345 350
Glu Leu Ala Ile Gln Val Gly Ala Glu Val Ile Glu Ser Ile Cys Asp
355 360 365
Gly Thr Pro Leu Ser Glu Val Leu Ala Asn Leu Ser Pro Arg Val Thr
370 375 380
Ile Gln Lys Glu Thr Gly Asn Asn Leu Leu Lys Phe Tyr Ala Leu Leu
385 390 395 400
Thr Val

<210> 42
<211> 2199
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2199)

<400> 42
atg ctg ttg cag ggc cct gtt ctt ctg ccc gcg tgt cct gct acg gtg 48
Met Leu Leu Gln Gly Pro Val Leu Leu Pro Ala Cys Pro Ala Thr Val
1 5 10 15

gca gcg aac gcc ccg tca ccg gcc aat tcc gat ttt aaa act caa ctg 96
Ala Ala Asn Ala Pro Ser Pro Ala Asn Ser Asp Phe Lys Thr Gln Leu
20 25 30

```
gca att ttt tgc tgc ctg gcc acc aat aat gag att ctt gaa aac gtc    144
Ala Ile Phe Cys Cys Leu Ala Thr Asn Asn Glu Ile Leu Glu Asn Val
         35                  40                  45

tcg cta gag gcc ttg gac agg gct atg caa aca gag acc acc ttt tac    192
Ser Leu Glu Ala Leu Asp Arg Ala Met Gln Thr Glu Thr Thr Phe Tyr
     50                  55                  60

gcg tgc cgt gcg ctc cgg cgt ctg gta ttg gga gaa gga cta tac cct    240
Ala Cys Arg Ala Leu Arg Arg Leu Val Leu Gly Glu Gly Leu Tyr Pro
 65                  70                  75                  80

ttt ata cac cgc caa ggt ggt atc gtt ggc aag acc ggt aac gag tac    288
Phe Ile His Arg Gln Gly Gly Ile Val Gly Lys Thr Gly Asn Glu Tyr
                 85                  90                  95

gcg gga ccg ggt tta ata att gat gat gct atc gga tgt aca ttt tcc    336
Ala Gly Pro Gly Leu Ile Ile Asp Asp Ala Ile Gly Cys Thr Phe Ser
            100                 105                 110

cat atc gag acg cat acg ttt tta ccc acc gtt ttc acc tac gaa ctg    384
His Ile Glu Thr His Thr Phe Leu Pro Thr Val Phe Thr Tyr Glu Leu
        115                 120                 125

tcc gac acg gtt ttg gtt caa agc gac gaa cga atc cta cgg tcg ctg    432
Ser Asp Thr Val Leu Val Gln Ser Asp Glu Arg Ile Leu Arg Ser Leu
    130                 135                 140

tac tgt tcg cca tta atg gta tgt ggc gta aat tat caa agc atg ttt    480
Tyr Cys Ser Pro Leu Met Val Cys Gly Val Asn Tyr Gln Ser Met Phe
145                 150                 155                 160

cga atc ctc tgt cgt tac ctt caa ata tgg gaa ttc gag gaa tgc ttt    528
Arg Ile Leu Cys Arg Tyr Leu Gln Ile Trp Glu Phe Glu Glu Cys Phe
                165                 170                 175

gct gcc ttt acg cgc aca ctg ccg gaa cat ctg atc ggc acg tgt tat    576
Ala Ala Phe Thr Arg Thr Leu Pro Glu His Leu Ile Gly Thr Cys Tyr
            180                 185                 190

cag aat tac ttt aaa ctt tta gag cca ttc aaa atc ctc acg cta gct    624
Gln Asn Tyr Phe Lys Leu Leu Glu Pro Phe Lys Ile Leu Thr Leu Ala
        195                 200                 205

cgg tgt ccg ccc ccg tgt gca aaa ctg cat ctg aat tac tta aaa ttt    672
Arg Cys Pro Pro Pro Cys Ala Lys Leu His Leu Asn Tyr Leu Lys Phe
    210                 215                 220

aac att ttg ggt ttt act tcg gat tgg att tcg cat ccc gaa ctg cat    720
Asn Ile Leu Gly Phe Thr Ser Asp Trp Ile Ser His Pro Glu Leu His
225                 230                 235                 240

cgc gtc caa acg gta att att cac aac atc gaa tcc aat ccg gta ctt    768
Arg Val Gln Thr Val Ile Ile His Asn Ile Glu Ser Asn Pro Val Leu
                245                 250                 255

tta aaa aac ctc tca aag caa aac aaa ttt cag gac atc aaa gtg gcg    816
Leu Lys Asn Leu Ser Lys Gln Asn Lys Phe Gln Asp Ile Lys Val Ala
            260                 265                 270

agc gaa tta atc ata gac tac caa aac ata gta aat caa tcg ctt gac    864
```

Ser Glu Leu Ile Ile Asp Tyr Gln Asn Ile Val Asn Gln Ser Leu Asp
275 280 285

gta aat cta cag gtg aaa att aac aaa aag gat ccc ggc aaa aaa ccg 912
Val Asn Leu Gln Val Lys Ile Asn Lys Lys Asp Pro Gly Lys Lys Pro
290 295 300

tac aag gta gtt gtt gtg acg cca aag tca acg tac tat cta acg ttt 960
Tyr Lys Val Val Val Val Thr Pro Lys Ser Thr Tyr Tyr Leu Thr Phe
305 310 315 320

cca ccc gag gta cca ata ttc cgc gtc gca atg tgc atg tct gtg gcg 1008
Pro Pro Glu Val Pro Ile Phe Arg Val Ala Met Cys Met Ser Val Ala
325 330 335

gaa cac gtg tgc cat tcg tgc gac agg ctg tat cct aac acc gag ttt 1056
Glu His Val Cys His Ser Cys Asp Arg Leu Tyr Pro Asn Thr Glu Phe
340 345 350

ctg ggg cca ggc gaa acg cct cgc gtc cta gag gcc atg ttt tcg cga 1104
Leu Gly Pro Gly Glu Thr Pro Arg Val Leu Glu Ala Met Phe Ser Arg
355 360 365

att caa tac gca cca aaa gat aga gat tac aat ttt att ttt aac gcc 1152
Ile Gln Tyr Ala Pro Lys Asp Arg Asp Tyr Asn Phe Ile Phe Asn Ala
370 375 380

gat caa aac cca gat cga tat gaa caa gcg cgt cac gac cac cag acc 1200
Asp Gln Asn Pro Asp Arg Tyr Glu Gln Ala Arg His Asp His Gln Thr
385 390 395 400

gaa cct ctt cca gac atg ttc gat cca gtt aag cac atg agt tta cac 1248
Glu Pro Leu Pro Asp Met Phe Asp Pro Val Lys His Met Ser Leu His
405 410 415

aac ttc aaa atc tca gtt ttc aat acc aat atg gta att aat acc aaa 1296
Asn Phe Lys Ile Ser Val Phe Asn Thr Asn Met Val Ile Asn Thr Lys
420 425 430

atc acg tgc tgg tct ctc gcc ggt acc ttc gag tca att atc gac atc 1344
Ile Thr Cys Trp Ser Leu Ala Gly Thr Phe Glu Ser Ile Ile Asp Ile
435 440 445

ccc agg ctc act aat aac ttt gta atg aaa aaa ttc tcc gtg aaa gaa 1392
Pro Arg Leu Thr Asn Asn Phe Val Met Lys Lys Phe Ser Val Lys Glu
450 455 460

ccg tca ttc acc gtg agt gtg ttc tac tcc gac aac ctg tgc aac ggg 1440
Pro Ser Phe Thr Val Ser Val Phe Tyr Ser Asp Asn Leu Cys Asn Gly
465 470 475 480

gct gca att aac gtt aac ata agc ggg gac atg ctg cac ttt atg ttc 1488
Ala Ala Ile Asn Val Asn Ile Ser Gly Asp Met Leu His Phe Met Phe
485 490 495

gct atg gga aac ctc agg tgc ttt tta ccg gtg aag cac att ttc ccg 1536
Ala Met Gly Asn Leu Arg Cys Phe Leu Pro Val Lys His Ile Phe Pro
500 505 510

gtt tcg att gcg aac tgg aac tct acg tta gac ctc cac ggg ctt gaa 1584
Val Ser Ile Ala Asn Trp Asn Ser Thr Leu Asp Leu His Gly Leu Glu

515 520 525

aac caa tac ata gtt aga cgg ggg cgg cga gac gtt ttc tgg acc act 1632
Asn Gln Tyr Ile Val Arg Arg Gly Arg Arg Asp Val Phe Trp Thr Thr
530 535 540

aat ttc cct tct gtg gtt tct agt aaa gac gga tgc aac gtt tcc tgg 1680
Asn Phe Pro Ser Val Val Ser Ser Lys Asp Gly Cys Asn Val Ser Trp
545 550 555 560

ttt aag gcg gca act gcg aca att tct aaa ata tac ggt agg ccc tta 1728
Phe Lys Ala Ala Thr Ala Thr Ile Ser Lys Ile Tyr Gly Arg Pro Leu
565 570 575

ctt aaa aaa ctt tcg gat gaa ttg aac cca ata tta agc gtt cca tat 1776
Leu Lys Lys Leu Ser Asp Glu Leu Asn Pro Ile Leu Ser Val Pro Tyr
580 585 590

gct aga ata gac cag gtg aaa aac aca ata ttc aca acg ctg gag aca 1824
Ala Arg Ile Asp Gln Val Lys Asn Thr Ile Phe Thr Thr Leu Glu Thr
595 600 605

aga aat aaa gcg caa ata caa act cta cac aag cgt ttt atc gag tgt 1872
Arg Asn Lys Ala Gln Ile Gln Thr Leu His Lys Arg Phe Ile Glu Cys
610 615 620

ttg gtg gaa tgt tgt tcg ttt ttg cgt ttg gat ctg gga gcc ctg aac 1920
Leu Val Glu Cys Cys Ser Phe Leu Arg Leu Asp Leu Gly Ala Leu Asn
625 630 635 640

cgc gcc gcc cgc ctt ggt acg ttt gac ttt tca aaa cgc atc att agt 1968
Arg Ala Ala Arg Leu Gly Thr Phe Asp Phe Ser Lys Arg Ile Ile Ser
645 650 655

cac acc aaa agc aaa cac gaa tgc gct att ctg gga tat aaa aaa tgc 2016
His Thr Lys Ser Lys His Glu Cys Ala Ile Leu Gly Tyr Lys Lys Cys
660 665 670

aac ctc att cca aag ata tac gtg cga agt aaa aaa ata cgg tta gac 2064
Asn Leu Ile Pro Lys Ile Tyr Val Arg Ser Lys Lys Ile Arg Leu Asp
675 680 685

gag tta gga cgc aac gcc aat ttt atg tct ttc ata gct acg acc ggt 2112
Glu Leu Gly Arg Asn Ala Asn Phe Met Ser Phe Ile Ala Thr Thr Gly
690 695 700

cac gcg ttt tca aac tta aaa ccg cag gta att cga cac act att aga 2160
His Ala Phe Ser Asn Leu Lys Pro Gln Val Ile Arg His Thr Ile Arg
705 710 715 720

cgt ttg gga ctt cac tgg aga cat aaa gca aaa ata taa 2199
Arg Leu Gly Leu His Trp Arg His Lys Ala Lys Ile
725 730

<210> 43
<211> 732
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 43

Met Leu Leu Gln Gly Pro Val Leu Leu Pro Ala Cys Pro Ala Thr Val
1 5 10 15
Ala Ala Asn Ala Pro Ser Pro Ala Asn Ser Asp Phe Lys Thr Gln Leu
20 25 30
Ala Ile Phe Cys Cys Leu Ala Thr Asn Asn Glu Ile Leu Glu Asn Val
35 40 45
Ser Leu Glu Ala Leu Asp Arg Ala Met Gln Thr Glu Thr Thr Phe Tyr
50 55 60
Ala Cys Arg Ala Leu Arg Arg Leu Val Leu Gly Glu Gly Leu Tyr Pro
65 70 75 80
Phe Ile His Arg Gln Gly Gly Ile Val Gly Lys Thr Gly Asn Glu Tyr
85 90 95
Ala Gly Pro Gly Leu Ile Ile Asp Asp Ala Ile Gly Cys Thr Phe Ser
100 105 110
His Ile Glu Thr His Thr Phe Leu Pro Thr Val Phe Thr Tyr Glu Leu
115 120 125
Ser Asp Thr Val Leu Val Gln Ser Asp Glu Arg Ile Leu Arg Ser Leu
130 135 140
Tyr Cys Ser Pro Leu Met Val Cys Gly Val Asn Tyr Gln Ser Met Phe
145 150 155 160
Arg Ile Leu Cys Arg Tyr Leu Gln Ile Trp Glu Phe Glu Glu Cys Phe
165 170 175
Ala Ala Phe Thr Arg Thr Leu Pro Glu His Leu Ile Gly Thr Cys Tyr
180 185 190
Gln Asn Tyr Phe Lys Leu Leu Glu Pro Phe Lys Ile Leu Thr Leu Ala
195 200 205
Arg Cys Pro Pro Pro Cys Ala Lys Leu His Leu Asn Tyr Leu Lys Phe
210 215 220
Asn Ile Leu Gly Phe Thr Ser Asp Trp Ile Ser His Pro Glu Leu His
225 230 235 240
Arg Val Gln Thr Val Ile Ile His Asn Ile Glu Ser Asn Pro Val Leu
245 250 255
Leu Lys Asn Leu Ser Lys Gln Asn Lys Phe Gln Asp Ile Lys Val Ala
260 265 270
Ser Glu Leu Ile Ile Asp Tyr Gln Asn Ile Val Asn Gln Ser Leu Asp
275 280 285
Val Asn Leu Gln Val Lys Ile Asn Lys Lys Asp Pro Gly Lys Lys Pro
290 295 300
Tyr Lys Val Val Val Val Thr Pro Lys Ser Thr Tyr Tyr Leu Thr Phe
305 310 315 320
Pro Pro Glu Val Pro Ile Phe Arg Val Ala Met Cys Met Ser Val Ala
325 330 335
Glu His Val Cys His Ser Cys Asp Arg Leu Tyr Pro Asn Thr Glu Phe
340 345 350
Leu Gly Pro Gly Glu Thr Pro Arg Val Leu Glu Ala Met Phe Ser Arg
355 360 365
Ile Gln Tyr Ala Pro Lys Asp Arg Asp Tyr Asn Phe Ile Phe Asn Ala
370 375 380
Asp Gln Asn Pro Asp Arg Tyr Glu Gln Ala Arg His Asp His Gln Thr
385 390 395 400
Glu Pro Leu Pro Asp Met Phe Asp Pro Val Lys His Met Ser Leu His
405 410 415
Asn Phe Lys Ile Ser Val Phe Asn Thr Asn Met Val Ile Asn Thr Lys
420 425 430
Ile Thr Cys Trp Ser Leu Ala Gly Thr Phe Glu Ser Ile Ile Asp Ile
435 440 445
Pro Arg Leu Thr Asn Asn Phe Val Met Lys Lys Phe Ser Val Lys Glu
450 455 460
Pro Ser Phe Thr Val Ser Val Phe Tyr Ser Asp Asn Leu Cys Asn Gly
465 470 475 480
Ala Ala Ile Asn Val Asn Ile Ser Gly Asp Met Leu His Phe Met Phe

485 490 495
Ala Met Gly Asn Leu Arg Cys Phe Leu Pro Val Lys His Ile Phe Pro
500 505 510
Val Ser Ile Ala Asn Trp Asn Ser Thr Leu Asp Leu His Gly Leu Glu
515 520 525
Asn Gln Tyr Ile Val Arg Arg Gly Arg Arg Asp Val Phe Trp Thr Thr
530 535 540
Asn Phe Pro Ser Val Val Ser Ser Lys Asp Gly Cys Asn Val Ser Trp
545 550 555 560
Phe Lys Ala Ala Thr Ala Thr Ile Ser Lys Ile Tyr Gly Arg Pro Leu
565 570 575
Leu Lys Lys Leu Ser Asp Glu Leu Asn Pro Ile Leu Ser Val Pro Tyr
580 585 590
Ala Arg Ile Asp Gln Val Lys Asn Thr Ile Phe Thr Thr Leu Glu Thr
595 600 605
Arg Asn Lys Ala Gln Ile Gln Thr Leu His Lys Arg Phe Ile Glu Cys
610 615 620
Leu Val Glu Cys Cys Ser Phe Leu Arg Leu Asp Leu Gly Ala Leu Asn
625 630 635 640
Arg Ala Ala Arg Leu Gly Thr Phe Asp Phe Ser Lys Arg Ile Ile Ser
645 650 655
His Thr Lys Ser Lys His Glu Cys Ala Ile Leu Gly Tyr Lys Lys Cys
660 665 670
Asn Leu Ile Pro Lys Ile Tyr Val Arg Ser Lys Lys Ile Arg Leu Asp
675 680 685
Glu Leu Gly Arg Asn Ala Asn Phe Met Ser Phe Ile Ala Thr Thr Gly
690 695 700
His Ala Phe Ser Asn Leu Lys Pro Gln Val Ile Arg His Thr Ile Arg
705 710 715 720
Arg Leu Gly Leu His Trp Arg His Lys Ala Lys Ile
725 730

<210> 44
<211> 4137
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(4137)

<400> 44
atg gag gcg gct ctg gag gta cga ccc ttt cca tat atg gcc acg gag 48
Met Glu Ala Ala Leu Glu Val Arg Pro Phe Pro Tyr Met Ala Thr Glu
1 5 10 15

gcc aat ctg ctt cgt caa atg aaa gag tcg gca gcc agc gga ctt ttt 96
Ala Asn Leu Leu Arg Gln Met Lys Glu Ser Ala Ala Ser Gly Leu Phe
20 25 30

aaa agt ttt caa ctg ctg ttg gga aaa gac gcc agg gag gga gga gtt 144
Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Gly Val
35 40 45

caa ttt gag ggt ctt ttg ggc gtg tat acc aac gta att cag ttt gtg 192
Gln Phe Glu Gly Leu Leu Gly Val Tyr Thr Asn Val Ile Gln Phe Val
50 55 60

aag ttt ctg gaa acg tcg cta gcg gtt gcg tgc gtt aac acc gag ttt 240

Lys Phe Leu Glu Thr Ser Leu Ala Val Ala Cys Val Asn Thr Glu Phe
65 70 75 80

aag gac ctc aaa cga atg acg gat gga aaa att cag ttt aag gta tct 288
Lys Asp Leu Lys Arg Met Thr Asp Gly Lys Ile Gln Phe Lys Val Ser
85 90 95

gta ccg acc atc gcg tat ggg gac ggc agg cgg ccc aca aaa caa aaa 336
Val Pro Thr Ile Ala Tyr Gly Asp Gly Arg Arg Pro Thr Lys Gln Lys
100 105 110

caa tac att atc atg aag gcc tgc aat aag cat cac att ggt gcc gag 384
Gln Tyr Ile Ile Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125

ata gag ctg tcg act gat gac atc gag ctg cta ttc att gac aga gaa 432
Ile Glu Leu Ser Thr Asp Asp Ile Glu Leu Leu Phe Ile Asp Arg Glu
130 135 140

acc cca ctc gat tac aca gaa tac gcc ggg gcc gta aag acg att acc 480
Thr Pro Leu Asp Tyr Thr Glu Tyr Ala Gly Ala Val Lys Thr Ile Thr
145 150 155 160

gcc tct ctc cag ttt ggc gtg gac gcg ctg gag agg ggc ctg gta gat 528
Ala Ser Leu Gln Phe Gly Val Asp Ala Leu Glu Arg Gly Leu Val Asp
165 170 175

acc gta ttg aat gtt aag ctt agg tcc gcc ccg ccg atg ttt att cta 576
Thr Val Leu Asn Val Lys Leu Arg Ser Ala Pro Pro Met Phe Ile Leu
180 185 190

aaa aca cta tca gac ccg gtc tac acc gaa cgg ggt cta aag aag gct 624
Lys Thr Leu Ser Asp Pro Val Tyr Thr Glu Arg Gly Leu Lys Lys Ala
195 200 205

gtt aag tca gac atg gtg tcc atg ttc aaa agc tac ctt atg gat aac 672
Val Lys Ser Asp Met Val Ser Met Phe Lys Ser Tyr Leu Met Asp Asn
210 215 220

tcg ttt ttc ctc gac aaa tca gac atc gcc gtc aag gga aag cag tac 720
Ser Phe Phe Leu Asp Lys Ser Asp Ile Ala Val Lys Gly Lys Gln Tyr
225 230 235 240

gtg ctg tcg gtt ctc tcc gac atg gtg ggg gcg gtg tgt cac gaa acg 768
Val Leu Ser Val Leu Ser Asp Met Val Gly Ala Val Cys His Glu Thr
245 250 255

gtt ttt aag ggg acg aat acg tat ctg tct gca tcg gga gag cca att 816
Val Phe Lys Gly Thr Asn Thr Tyr Leu Ser Ala Ser Gly Glu Pro Ile
260 265 270

gcc gga gtc atg gag acc acg gaa aac gta atg cga aaa ctg tta aac 864
Ala Gly Val Met Glu Thr Thr Glu Asn Val Met Arg Lys Leu Leu Asn
275 280 285

atg cta ggt cag gtt gac ggg ggc atg tcc ggt ccg gcg tct tac gcc 912
Met Leu Gly Gln Val Asp Gly Gly Met Ser Gly Pro Ala Ser Tyr Ala
290 295 300

aat tac gtt gtc agg ggc gaa aat ctc gta acc gcc gtg acg tac ggt 960
Asn Tyr Val Val Arg Gly Glu Asn Leu Val Thr Ala Val Thr Tyr Gly

305 310 315 320

cgc gtc atg cga acg ttt gac cag ttc atg aaa cgc ata gtg gac cgt 1008
Arg Val Met Arg Thr Phe Asp Gln Phe Met Lys Arg Ile Val Asp Arg
325 330 335

ccc aac gcg cag ccc agc gtc gac gat gat cgg gac gcg gtg gcg gac 1056
Pro Asn Ala Gln Pro Ser Val Asp Asp Asp Arg Asp Ala Val Ala Asp
340 345 350

ggg cag gac tcc ctc gcc aaa aca ccg atc gcg gcg gct gtt atc cag 1104
Gly Gln Asp Ser Leu Ala Lys Thr Pro Ile Ala Ala Ala Val Ile Gln
355 360 365

att gga gat aaa ctg gtg gct ttg gaa agt tta cag cga atg tac aac 1152
Ile Gly Asp Lys Leu Val Ala Leu Glu Ser Leu Gln Arg Met Tyr Asn
370 375 380

gag acc cag ttt ccc ttc ccg tta aac agg cgt atg cac tat act tat 1200
Glu Thr Gln Phe Pro Phe Pro Leu Asn Arg Arg Met His Tyr Thr Tyr
385 390 395 400

ttt ttc ccc att ggg ttg cac atg ccc cgt cct cag tat tcc acg tcg 1248
Phe Phe Pro Ile Gly Leu His Met Pro Arg Pro Gln Tyr Ser Thr Ser
405 410 415

gcc aca att aag ggg gtc gag cac cca gcg gaa caa tcc gtc gaa acg 1296
Ala Thr Ile Lys Gly Val Glu His Pro Ala Glu Gln Ser Val Glu Thr
420 425 430

tgg att gta aat aaa aac aac gta ctg ctg agt ttc aat tat caa aac 1344
Trp Ile Val Asn Lys Asn Asn Val Leu Leu Ser Phe Asn Tyr Gln Asn
435 440 445

gcg ctc aag tcc atc tgt cat ccg cgc atg cac aac ccg atg ccg tgt 1392
Ala Leu Lys Ser Ile Cys His Pro Arg Met His Asn Pro Met Pro Cys
450 455 460

ggg caa gct ttg ggt cag gcg ttt cca gat ccc gga cac gtt cac agg 1440
Gly Gln Ala Leu Gly Gln Ala Phe Pro Asp Pro Gly His Val His Arg
465 470 475 480

tac gga caa aga tct gag cac ccc ccg aac atg aac cta tac ggg ctg 1488
Tyr Gly Gln Arg Ser Glu His Pro Pro Asn Met Asn Leu Tyr Gly Leu
485 490 495

gtg tac aac tat tac cag ggc aaa aac gtg gca cac gtt ccg gat gtc 1536
Val Tyr Asn Tyr Tyr Gln Gly Lys Asn Val Ala His Val Pro Asp Val
500 505 510

gcc cta aag gcg acg atg acc aca gat gag ttg ctg cac cca acc tcg 1584
Ala Leu Lys Ala Thr Met Thr Thr Asp Glu Leu Leu His Pro Thr Ser
515 520 525

cac gag acg ctt cgt ttg gag gtt cac ccg atg ttt gat ttt ttc gtt 1632
His Glu Thr Leu Arg Leu Glu Val His Pro Met Phe Asp Phe Phe Val
530 535 540

cat cag cag cct ggt gcg caa gcc gcg tat agg gcg acc cac agg act 1680
His Gln Gln Pro Gly Ala Gln Ala Ala Tyr Arg Ala Thr His Arg Thr
545 550 555 560

atg gtg ggt aac att cca caa ccc ctg gcg ccc aac gag ttt caa aac 1728
Met Val Gly Asn Ile Pro Gln Pro Leu Ala Pro Asn Glu Phe Gln Asn
565 570 575

agc aga ggc ctg cag ttt gac aga gcg gcg gcc gtg gct cac gtg ctg 1776
Ser Arg Gly Leu Gln Phe Asp Arg Ala Ala Ala Val Ala His Val Leu
580 585 590

gac cag tca acg atg gaa att atc caa gat acg gcg ttt gac acg tcg 1824
Asp Gln Ser Thr Met Glu Ile Ile Gln Asp Thr Ala Phe Asp Thr Ser
595 600 605

tac cca cta ctc tgt tat gtc atc gaa tgc ctc att cac gga cag gaa 1872
Tyr Pro Leu Leu Cys Tyr Val Ile Glu Cys Leu Ile His Gly Gln Glu
610 615 620

gac aaa ttt ttg att aat tct cct tta att gca tta acc att gaa act 1920
Asp Lys Phe Leu Ile Asn Ser Pro Leu Ile Ala Leu Thr Ile Glu Thr
625 630 635 640

tac tgg aac aat gcc gga aaa ctg gcg ttt att aac agc ttc cct atg 1968
Tyr Trp Asn Asn Ala Gly Lys Leu Ala Phe Ile Asn Ser Phe Pro Met
645 650 655

ctg cga ttt atc tgc gtt cac ctg ggc aac ggt agt att tct aag gac 2016
Leu Arg Phe Ile Cys Val His Leu Gly Asn Gly Ser Ile Ser Lys Asp
660 665 670

gtg tac gcc cat tac cga aaa gtt ttt ggc gaa ctc gtg gtt ttg cag 2064
Val Tyr Ala His Tyr Arg Lys Val Phe Gly Glu Leu Val Val Leu Gln
675 680 685

cag gcg ctc tcg aaa atc gcg ggc cac gag gtg gtg ggg cgc agg ccc 2112
Gln Ala Leu Ser Lys Ile Ala Gly His Glu Val Val Gly Arg Arg Pro
690 695 700

gcg tcc gag ctg att aac tgt ctt cag gac ccc aat ctt ttg ccg ccc 2160
Ala Ser Glu Leu Ile Asn Cys Leu Gln Asp Pro Asn Leu Leu Pro Pro
705 710 715 720

ttt gct tac aat gac gtt ttt acc aac ctg ctc agg cag tcc tcg cgg 2208
Phe Ala Tyr Asn Asp Val Phe Thr Asn Leu Leu Arg Gln Ser Ser Arg
725 730 735

cac ccc atg gta ctc ata ggc gac gag ggg tac gaa acg gaa aat gac 2256
His Pro Met Val Leu Ile Gly Asp Glu Gly Tyr Glu Thr Glu Asn Asp
740 745 750

agg gat acg tac atc aac gtc aga gga aaa atg gag gac cta gtc ggt 2304
Arg Asp Thr Tyr Ile Asn Val Arg Gly Lys Met Glu Asp Leu Val Gly
755 760 765

gac atg gtt aac att tac gag acc aga aac aac gcg gat cat gac ggc 2352
Asp Met Val Asn Ile Tyr Glu Thr Arg Asn Asn Ala Asp His Asp Gly
770 775 780

cgc cac gtc ctt gac gtc ggt ccc ttt aat gaa aac gaa cag cac atg 2400
Arg His Val Leu Asp Val Gly Pro Phe Asn Glu Asn Glu Gln His Met
785 790 795 800

```
gct gtg ctg gaa aag ctt ttt tat tac gtg gtt ctg cca gcc tgt acc     2448
Ala Val Leu Glu Lys Leu Phe Tyr Tyr Val Val Leu Pro Ala Cys Thr
            805                 810                 815

aac ggt cac gtc tgc ggc atg ggc gtc gat ttt gac aac gtg gcc ctg     2496
Asn Gly His Val Cys Gly Met Gly Val Asp Phe Asp Asn Val Ala Leu
        820                 825                 830

gcc ctg acg tac aac ggc cca gtg ttt gct gac gtc gtg aac ccc gac     2544
Ala Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Pro Asp
    835                 840                 845

gat gag att ttg gac cac ctg gag aac ggg acg ctc cgc gag atg ctc     2592
Asp Glu Ile Leu Asp His Leu Glu Asn Gly Thr Leu Arg Glu Met Leu
850                 855                 860

gag gct tcg gat ata cac ccc acc gtt gac atg att cga act ctt tgc     2640
Glu Ala Ser Asp Ile His Pro Thr Val Asp Met Ile Arg Thr Leu Cys
865                 870                 875                 880

acg tcg ttt ctc acc tgc ccg ttt gtt acc cag gcc tcc cgt gtt gtg     2688
Thr Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ser Arg Val Val
            885                 890                 895

act cag cgg gac ccc gcg caa ctg ttg acc act cac gac gac ggg aga     2736
Thr Gln Arg Asp Pro Ala Gln Leu Leu Thr Thr His Asp Asp Gly Arg
        900                 905                 910

tac gtg agc cag act gtc ctc gtt aac ggg ttc gcg gcg ttt gct atc     2784
Tyr Val Ser Gln Thr Val Leu Val Asn Gly Phe Ala Ala Phe Ala Ile
    915                 920                 925

gca gat agg tct cgt gac gtt gcc gag acc atg ttt tac ccg gtg ccg     2832
Ala Asp Arg Ser Arg Asp Val Ala Glu Thr Met Phe Tyr Pro Val Pro
930                 935                 940

ttc acc aag ctg tac agc gat ccc ctg gtg gcg gcc acg ctc cac ccg     2880
Phe Thr Lys Leu Tyr Ser Asp Pro Leu Val Ala Ala Thr Leu His Pro
945                 950                 955                 960

ctg gtc gca aat tac gtg acg cgc ctg ccg gcc cag cgc gtg ccg gtc     2928
Leu Val Ala Asn Tyr Val Thr Arg Leu Pro Ala Gln Arg Val Pro Val
            965                 970                 975

gcg ttt aac gtc ccc ccg gcc ctc atg gcc gag tac gag gag tgg cac     2976
Ala Phe Asn Val Pro Pro Ala Leu Met Ala Glu Tyr Glu Glu Trp His
        980                 985                 990

aag tct cca atg ctg gcc tac gct aac acc tgc ccg atg acg ccc acg     3024
Lys Ser Pro Met Leu Ala Tyr Ala Asn Thr Cys Pro Met Thr Pro Thr
    995                1000                1005

tcg ttg agc acc ctg gcg agc atg cac atg aag ctg tcc gcg ccg ggg     3072
Ser Leu Ser Thr Leu Ala Ser Met His Met Lys Leu Ser Ala Pro Gly
1010                1015                1020

ttc atc tgc cac gca aag cac aag att cac ccg ggc ttt gcg atg acc     3120
Phe Ile Cys His Ala Lys His Lys Ile His Pro Gly Phe Ala Met Thr
1025                1030                1035                1040

gcc gtc cga acc gat gag gtg ttg gcg gag aac ttg cta ttt agt gcc     3168
```

Ala Val Arg Thr Asp Glu Val Leu Ala Glu Asn Leu Leu Phe Ser Ala
1045 1050 1055

agg gcc tcg acg tcc atg ttt tta ggg cag cca tcg gtt atg cgt cgg 3216
Arg Ala Ser Thr Ser Met Phe Leu Gly Gln Pro Ser Val Met Arg Arg
1060 1065 1070

gaa gtc agg gcg gac gca gtc acg ttt gag gtg aat cat gag ttg gca 3264
Glu Val Arg Ala Asp Ala Val Thr Phe Glu Val Asn His Glu Leu Ala
1075 1080 1085

tcg ctg gac atg gcg ctc ggt tat tct tcc acc atc acg ccc gcc cac 3312
Ser Leu Asp Met Ala Leu Gly Tyr Ser Ser Thr Ile Thr Pro Ala His
1090 1095 1100

gtt gcg gcg att acc tcg gac atg ggc gtt cac tgt cag gac atg ttt 3360
Val Ala Ala Ile Thr Ser Asp Met Gly Val His Cys Gln Asp Met Phe
1105 1110 1115 1120

ctc atg ttt ccc ggg gac tcg tac cag gac agg acc ctc aac gac tac 3408
Leu Met Phe Pro Gly Asp Ser Tyr Gln Asp Arg Thr Leu Asn Asp Tyr
1125 1130 1135

gtt aaa caa aaa gcc gga tgc caa cga ttc ggt ggt cct ggc cag att 3456
Val Lys Gln Lys Ala Gly Cys Gln Arg Phe Gly Gly Pro Gly Gln Ile
1140 1145 1150

cgt gag ccc gtc gct tac gtt gcg ggg gtg ccg cac tcg gac aac ata 3504
Arg Glu Pro Val Ala Tyr Val Ala Gly Val Pro His Ser Asp Asn Ile
1155 1160 1165

ccg ggt ctc agc cac gga cag ctg gcc acg tgt gag att gtt ttg acg 3552
Pro Gly Leu Ser His Gly Gln Leu Ala Thr Cys Glu Ile Val Leu Thr
1170 1175 1180

ccc gtt act gca gac gtt acc tat ttt caa acc ccc aac agt ccc cgg 3600
Pro Val Thr Ala Asp Val Thr Tyr Phe Gln Thr Pro Asn Ser Pro Arg
1185 1190 1195 1200

gga cgg gca tcc tgc gtg atc tcg tgt gac gcg tac aac aac gaa agc 3648
Gly Arg Ala Ser Cys Val Ile Ser Cys Asp Ala Tyr Asn Asn Glu Ser
1205 1210 1215

gcg gaa cgt ttg ctc ttt gac cac tcc atc ccg gat tct gcc tac gaa 3696
Ala Glu Arg Leu Leu Phe Asp His Ser Ile Pro Asp Ser Ala Tyr Glu
1220 1225 1230

tac cgc act acg gtt aac cca tgg gcg tcg cag cag ggc tcc ctc gga 3744
Tyr Arg Thr Thr Val Asn Pro Trp Ala Ser Gln Gln Gly Ser Leu Gly
1235 1240 1245

gac gtg ctg tac aac tca acc tcg cgc cag gtc gca gtg cca ggg atg 3792
Asp Val Leu Tyr Asn Ser Thr Ser Arg Gln Val Ala Val Pro Gly Met
1250 1255 1260

tac agt ccg tgt cgc cag ttt ttc cac aag gac gct att ttg cgt aac 3840
Tyr Ser Pro Cys Arg Gln Phe Phe His Lys Asp Ala Ile Leu Arg Asn
1265 1270 1275 1280

aat cgg ggc ctg aac aca cta gtc acg gaa tac gcg gcc cgc ctc acg 3888
Asn Arg Gly Leu Asn Thr Leu Val Thr Glu Tyr Ala Ala Arg Leu Thr

1285 1290 1295

gga acg ccg gcg acc agc gcg acg gac ctg cag tac gtg gtg gtc aac 3936
Gly Thr Pro Ala Thr Ser Ala Thr Asp Leu Gln Tyr Val Val Val Asn
1300 1305 1310

gga acg gat gtg ttt cta gaa caa ccg tgc cag ttt cta caa gaa gcg 3984
Gly Thr Asp Val Phe Leu Glu Gln Pro Cys Gln Phe Leu Gln Glu Ala
1315 1320 1325

ttt ccc acg ctc gcc gcc agt cac agg tct ctg ctg gac gaa tat atg 4032
Phe Pro Thr Leu Ala Ala Ser His Arg Ser Leu Leu Asp Glu Tyr Met
1330 1335 1340

tcg aat aag ctc acg cac gcc cct gtg cac atg gga cat tat atg att 4080
Ser Asn Lys Leu Thr His Ala Pro Val His Met Gly His Tyr Met Ile
1345 1350 1355 1360

gag gaa gtg gcc cct atg aaa aga cta tta aag atc gga aac aag gtc 4128
Glu Glu Val Ala Pro Met Lys Arg Leu Leu Lys Ile Gly Asn Lys Val
1365 1370 1375

gcc tat tag 4137
Ala Tyr

<210> 45
<211> 1378
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 45
Met Glu Ala Ala Leu Glu Val Arg Pro Phe Pro Tyr Met Ala Thr Glu
1 5 10 15
Ala Asn Leu Leu Arg Gln Met Lys Glu Ser Ala Ala Ser Gly Leu Phe
20 25 30
Lys Ser Phe Gln Leu Leu Leu Gly Lys Asp Ala Arg Glu Gly Gly Val
35 40 45
Gln Phe Glu Gly Leu Leu Gly Val Tyr Thr Asn Val Ile Gln Phe Val
50 55 60
Lys Phe Leu Glu Thr Ser Leu Ala Val Ala Cys Val Asn Thr Glu Phe
65 70 75 80
Lys Asp Leu Lys Arg Met Thr Asp Gly Lys Ile Gln Phe Lys Val Ser
85 90 95
Val Pro Thr Ile Ala Tyr Gly Asp Gly Arg Arg Pro Thr Lys Gln Lys
100 105 110
Gln Tyr Ile Ile Met Lys Ala Cys Asn Lys His His Ile Gly Ala Glu
115 120 125
Ile Glu Leu Ser Thr Asp Asp Ile Glu Leu Leu Phe Ile Asp Arg Glu
130 135 140
Thr Pro Leu Asp Tyr Thr Glu Tyr Ala Gly Ala Val Lys Thr Ile Thr
145 150 155 160
Ala Ser Leu Gln Phe Gly Val Asp Ala Leu Glu Arg Gly Leu Val Asp
165 170 175
Thr Val Leu Asn Val Lys Leu Arg Ser Ala Pro Pro Met Phe Ile Leu
180 185 190
Lys Thr Leu Ser Asp Pro Val Tyr Thr Glu Arg Gly Leu Lys Lys Ala
195 200 205
Val Lys Ser Asp Met Val Ser Met Phe Lys Ser Tyr Leu Met Asp Asn
210 215 220
Ser Phe Phe Leu Asp Lys Ser Asp Ile Ala Val Lys Gly Lys Gln Tyr

225 230 235 240
Val Leu Ser Val Leu Ser Asp Met Val Gly Ala Val Cys His Glu Thr
245 250 255
Val Phe Lys Gly Thr Asn Thr Tyr Leu Ser Ala Ser Gly Glu Pro Ile
260 265 270
Ala Gly Val Met Glu Thr Thr Glu Asn Val Met Arg Lys Leu Leu Asn
275 280 285
Met Leu Gly Gln Val Asp Gly Gly Met Ser Gly Pro Ala Ser Tyr Ala
290 295 300
Asn Tyr Val Val Arg Gly Glu Asn Leu Val Thr Ala Val Thr Tyr Gly
305 310 315 320
Arg Val Met Arg Thr Phe Asp Gln Phe Met Lys Arg Ile Val Asp Arg
325 330 335
Pro Asn Ala Gln Pro Ser Val Asp Asp Asp Arg Asp Ala Val Ala Asp
340 345 350
Gly Gln Asp Ser Leu Ala Lys Thr Pro Ile Ala Ala Ala Val Ile Gln
355 360 365
Ile Gly Asp Lys Leu Val Ala Leu Glu Ser Leu Gln Arg Met Tyr Asn
370 375 380
Glu Thr Gln Phe Pro Phe Pro Leu Asn Arg Arg Met His Tyr Thr Tyr
385 390 395 400
Phe Phe Pro Ile Gly Leu His Met Pro Arg Pro Gln Tyr Ser Thr Ser
405 410 415
Ala Thr Ile Lys Gly Val Glu His Pro Ala Glu Gln Ser Val Glu Thr
420 425 430
Trp Ile Val Asn Lys Asn Asn Val Leu Leu Ser Phe Asn Tyr Gln Asn
435 440 445
Ala Leu Lys Ser Ile Cys His Pro Arg Met His Asn Pro Met Pro Cys
450 455 460
Gly Gln Ala Leu Gly Gln Ala Phe Pro Asp Pro Gly His Val His Arg
465 470 475 480
Tyr Gly Gln Arg Ser Glu His Pro Pro Asn Met Asn Leu Tyr Gly Leu
485 490 495
Val Tyr Asn Tyr Tyr Gln Gly Lys Asn Val Ala His Val Pro Asp Val
500 505 510
Ala Leu Lys Ala Thr Met Thr Thr Asp Glu Leu Leu His Pro Thr Ser
515 520 525
His Glu Thr Leu Arg Leu Glu Val His Pro Met Phe Asp Phe Phe Val
530 535 540
His Gln Gln Pro Gly Ala Gln Ala Ala Tyr Arg Ala Thr His Arg Thr
545 550 555 560
Met Val Gly Asn Ile Pro Gln Pro Leu Ala Pro Asn Glu Phe Gln Asn
565 570 575
Ser Arg Gly Leu Gln Phe Asp Arg Ala Ala Ala Val Ala His Val Leu
580 585 590
Asp Gln Ser Thr Met Glu Ile Ile Gln Asp Thr Ala Phe Asp Thr Ser
595 600 605
Tyr Pro Leu Leu Cys Tyr Val Ile Glu Cys Leu Ile His Gly Gln Glu
610 615 620
Asp Lys Phe Leu Ile Asn Ser Pro Leu Ile Ala Leu Thr Ile Glu Thr
625 630 635 640
Tyr Trp Asn Asn Ala Gly Lys Leu Ala Phe Ile Asn Ser Phe Pro Met
645 650 655
Leu Arg Phe Ile Cys Val His Leu Gly Asn Gly Ser Ile Ser Lys Asp
660 665 670
Val Tyr Ala His Tyr Arg Lys Val Phe Gly Glu Leu Val Val Leu Gln
675 680 685
Gln Ala Leu Ser Lys Ile Ala Gly His Glu Val Val Gly Arg Arg Pro
690 695 700
Ala Ser Glu Leu Ile Asn Cys Leu Gln Asp Pro Asn Leu Leu Pro Pro
705 710 715 720

Phe Ala Tyr Asn Asp Val Phe Thr Asn Leu Leu Arg Gln Ser Ser Arg
725 730 735
His Pro Met Val Leu Ile Gly Asp Glu Gly Tyr Glu Thr Glu Asn Asp
740 745 750
Arg Asp Thr Tyr Ile Asn Val Arg Gly Lys Met Glu Asp Leu Val Gly
755 760 765
Asp Met Val Asn Ile Tyr Glu Thr Arg Asn Asn Ala Asp His Asp Gly
770 775 780
Arg His Val Leu Asp Val Gly Pro Phe Asn Glu Asn Glu Gln His Met
785 790 795 800
Ala Val Leu Glu Lys Leu Phe Tyr Tyr Val Val Leu Pro Ala Cys Thr
805 810 815
Asn Gly His Val Cys Gly Met Gly Val Asp Phe Asp Asn Val Ala Leu
820 825 830
Ala Leu Thr Tyr Asn Gly Pro Val Phe Ala Asp Val Val Asn Pro Asp
835 840 845
Asp Glu Ile Leu Asp His Leu Glu Asn Gly Thr Leu Arg Glu Met Leu
850 855 860
Glu Ala Ser Asp Ile His Pro Thr Val Asp Met Ile Arg Thr Leu Cys
865 870 875 880
Thr Ser Phe Leu Thr Cys Pro Phe Val Thr Gln Ala Ser Arg Val Val
885 890 895
Thr Gln Arg Asp Pro Ala Gln Leu Leu Thr Thr His Asp Asp Gly Arg
900 905 910
Tyr Val Ser Gln Thr Val Leu Val Asn Gly Phe Ala Ala Phe Ala Ile
915 920 925
Ala Asp Arg Ser Arg Asp Val Ala Glu Thr Met Phe Tyr Pro Val Pro
930 935 940
Phe Thr Lys Leu Tyr Ser Asp Pro Leu Val Ala Ala Thr Leu His Pro
945 950 955 960
Leu Val Ala Asn Tyr Val Thr Arg Leu Pro Ala Gln Arg Val Pro Val
965 970 975
Ala Phe Asn Val Pro Pro Ala Leu Met Ala Glu Tyr Glu Glu Trp His
980 985 990
Lys Ser Pro Met Leu Ala Tyr Ala Asn Thr Cys Pro Met Thr Pro Thr
995 1000 1005
Ser Leu Ser Thr Leu Ala Ser Met His Met Lys Leu Ser Ala Pro Gly
1010 1015 1020
Phe Ile Cys His Ala Lys His Lys Ile His Pro Gly Phe Ala Met Thr
1025 1030 1035 1040
Ala Val Arg Thr Asp Glu Val Leu Ala Glu Asn Leu Leu Phe Ser Ala
1045 1050 1055
Arg Ala Ser Thr Ser Met Phe Leu Gly Gln Pro Ser Val Met Arg Arg
1060 1065 1070
Glu Val Arg Ala Asp Ala Val Thr Phe Glu Val Asn His Glu Leu Ala
1075 1080 1085
Ser Leu Asp Met Ala Leu Gly Tyr Ser Ser Thr Ile Thr Pro Ala His
1090 1095 1100
Val Ala Ala Ile Thr Ser Asp Met Gly Val His Cys Gln Asp Met Phe
1105 1110 1115 1120
Leu Met Phe Pro Gly Asp Ser Tyr Gln Asp Arg Thr Leu Asn Asp Tyr
1125 1130 1135
Val Lys Gln Lys Ala Gly Cys Gln Arg Phe Gly Gly Pro Gly Gln Ile
1140 1145 1150
Arg Glu Pro Val Ala Tyr Val Ala Gly Val Pro His Ser Asp Asn Ile
1155 1160 1165
Pro Gly Leu Ser His Gly Gln Leu Ala Thr Cys Glu Ile Val Leu Thr
1170 1175 1180
Pro Val Thr Ala Asp Val Thr Tyr Phe Gln Thr Pro Asn Ser Pro Arg
1185 1190 1195 1200
Gly Arg Ala Ser Cys Val Ile Ser Cys Asp Ala Tyr Asn Asn Glu Ser

1205 1210 1215
Ala Glu Arg Leu Leu Phe Asp His Ser Ile Pro Asp Ser Ala Tyr Glu
1220 1225 1230
Tyr Arg Thr Thr Val Asn Pro Trp Ala Ser Gln Gln Gly Ser Leu Gly
1235 1240 1245
Asp Val Leu Tyr Asn Ser Thr Ser Arg Gln Val Ala Val Pro Gly Met
1250 1255 1260
Tyr Ser Pro Cys Arg Gln Phe Phe His Lys Asp Ala Ile Leu Arg Asn
1265 1270 1275 1280
Asn Arg Gly Leu Asn Thr Leu Val Thr Glu Tyr Ala Ala Arg Leu Thr
1285 1290 1295
Gly Thr Pro Ala Thr Ser Ala Thr Asp Leu Gln Tyr Val Val Val Asn
1300 1305 1310
Gly Thr Asp Val Phe Leu Glu Gln Pro Cys Gln Phe Leu Gln Glu Ala
1315 1320 1325
Phe Pro Thr Leu Ala Ala Ser His Arg Ser Leu Leu Asp Glu Tyr Met
1330 1335 1340
Ser Asn Lys Leu Thr His Ala Pro Val His Met Gly His Tyr Met Ile
1345 1350 1355 1360
Glu Glu Val Ala Pro Met Lys Arg Leu Leu Lys Ile Gly Asn Lys Val
1365 1370 1375
Ala Tyr

<210> 46
<211> 918
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(918)

<400> 46
atg gcc ctc gat aag agc atc gtt gtc tcg gtg acg tct aga tta ttc 48
Met Ala Leu Asp Lys Ser Ile Val Val Ser Val Thr Ser Arg Leu Phe
1 5 10 15

gcc gac gag ata gca aat ctt cag tca aag ata gga tgc att ttg cct 96
Ala Asp Glu Ile Ala Asn Leu Gln Ser Lys Ile Gly Cys Ile Leu Pro
20 25 30

ctc aga gac gcc cac cgt ctg cag aat ata cag gcg ctg ggt ctg ggg 144
Leu Arg Asp Ala His Arg Leu Gln Asn Ile Gln Ala Leu Gly Leu Gly
35 40 45

aac ctg tgc tct agg gat tcc gcg gtg gat ttt att cag gca tat cac 192
Asn Leu Cys Ser Arg Asp Ser Ala Val Asp Phe Ile Gln Ala Tyr His
50 55 60

tat ttg gac aaa tgc act ctc gcc gtg ttg gaa gag gtc ggt ccc aac 240
Tyr Leu Asp Lys Cys Thr Leu Ala Val Leu Glu Glu Val Gly Pro Asn
65 70 75 80

agt tta cgg cta acg cgc att gat ccc atg gac aat tat caa ata aaa 288
Ser Leu Arg Leu Thr Arg Ile Asp Pro Met Asp Asn Tyr Gln Ile Lys
85 90 95

aac gcg tac caa ccg gcc ttc cat tgg gat aac tac tca gaa ttg gta 336
Asn Ala Tyr Gln Pro Ala Phe His Trp Asp Asn Tyr Ser Glu Leu Val

100 105 110

gtt ata cca ccg gtc ttt ggg cgc aaa gat gcg acc gtc tca ctg gag 384
Val Ile Pro Pro Val Phe Gly Arg Lys Asp Ala Thr Val Ser Leu Glu
115 120 125

tct aac ggg ttt gat gtg gtt ttc cct gcc gtg gtg cca gaa cca ctg 432
Ser Asn Gly Phe Asp Val Val Phe Pro Ala Val Val Pro Glu Pro Leu
130 135 140

gct caa aca gtg ctt cag aag ctg ctg ctg tat aac ata tac tac aga 480
Ala Gln Thr Val Leu Gln Lys Leu Leu Leu Tyr Asn Ile Tyr Tyr Arg
145 150 155 160

gtg gcg gag acg acg ccc acc gac gtc aac cta gcc gag gtg acg ctg 528
Val Ala Glu Thr Thr Pro Thr Asp Val Asn Leu Ala Glu Val Thr Leu
165 170 175

tac acg acc aat atc act tac atg ggt cgc aac tac gcc ctg gac gtg 576
Tyr Thr Thr Asn Ile Thr Tyr Met Gly Arg Asn Tyr Ala Leu Asp Val
180 185 190

gac ccc gtt ggg tcg agc tca gct atg cgg atg ctg gac gac ctg tcc 624
Asp Pro Val Gly Ser Ser Ser Ala Met Arg Met Leu Asp Asp Leu Ser
195 200 205

att tac ctg tgc gtt ttg tcc gcg tta att ccg cgc ggg tgc gta agg 672
Ile Tyr Leu Cys Val Leu Ser Ala Leu Ile Pro Arg Gly Cys Val Arg
210 215 220

cta ctg acc tca ttg gtg cgc cac aac aaa cac gaa tta gtc gag att 720
Leu Leu Thr Ser Leu Val Arg His Asn Lys His Glu Leu Val Glu Ile
225 230 235 240

ttc gag ggg gtg gtg cca cct gag gta cag gcc ctg gat ctc aac aac 768
Phe Glu Gly Val Val Pro Pro Glu Val Gln Ala Leu Asp Leu Asn Asn
245 250 255

gta agc gtg gcc gac gac ata acg cgc atg ggt gcc ctc ata acc tat 816
Val Ser Val Ala Asp Asp Ile Thr Arg Met Gly Ala Leu Ile Thr Tyr
260 265 270

cta cga agt ctc agt tct ata ttt aat ctg ggc cgc aga ttt cac gtt 864
Leu Arg Ser Leu Ser Ser Ile Phe Asn Leu Gly Arg Arg Phe His Val
275 280 285

tac gcg ttc tca tcg gac acg aat acc gct tcc tgt tgg tgt gca tat 912
Tyr Ala Phe Ser Ser Asp Thr Asn Thr Ala Ser Cys Trp Cys Ala Tyr
290 295 300

aac tag 918
Asn
305

<210> 47
<211> 305
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 47

Met Ala Leu Asp Lys Ser Ile Val Val Ser Val Thr Ser Arg Leu Phe
1 5 10 15
Ala Asp Glu Ile Ala Asn Leu Gln Ser Lys Ile Gly Cys Ile Leu Pro
20 25 30
Leu Arg Asp Ala His Arg Leu Gln Asn Ile Gln Ala Leu Gly Leu Gly
35 40 45
Asn Leu Cys Ser Arg Asp Ser Ala Val Asp Phe Ile Gln Ala Tyr His
50 55 60
Tyr Leu Asp Lys Cys Thr Leu Ala Val Leu Glu Glu Val Gly Pro Asn
65 70 75 80
Ser Leu Arg Leu Thr Arg Ile Asp Pro Met Asp Asn Tyr Gln Ile Lys
85 90 95
Asn Ala Tyr Gln Pro Ala Phe His Trp Asp Asn Tyr Ser Glu Leu Val
100 105 110
Val Ile Pro Pro Val Phe Gly Arg Lys Asp Ala Thr Val Ser Leu Glu
115 120 125
Ser Asn Gly Phe Asp Val Val Phe Pro Ala Val Val Pro Glu Pro Leu
130 135 140
Ala Gln Thr Val Leu Gln Lys Leu Leu Leu Tyr Asn Ile Tyr Tyr Arg
145 150 155 160
Val Ala Glu Thr Thr Pro Thr Asp Val Asn Leu Ala Glu Val Thr Leu
165 170 175
Tyr Thr Thr Asn Ile Thr Tyr Met Gly Arg Asn Tyr Ala Leu Asp Val
180 185 190
Asp Pro Val Gly Ser Ser Ser Ala Met Arg Met Leu Asp Asp Leu Ser
195 200 205
Ile Tyr Leu Cys Val Leu Ser Ala Leu Ile Pro Arg Gly Cys Val Arg
210 215 220
Leu Leu Thr Ser Leu Val Arg His Asn Lys His Glu Leu Val Glu Ile
225 230 235 240
Phe Glu Gly Val Val Pro Pro Glu Val Gln Ala Leu Asp Leu Asn Asn
245 250 255
Val Ser Val Ala Asp Asp Ile Thr Arg Met Gly Ala Leu Ile Thr Tyr
260 265 270
Leu Arg Ser Leu Ser Ser Ile Phe Asn Leu Gly Arg Arg Phe His Val
275 280 285
Tyr Ala Phe Ser Ser Asp Thr Asn Thr Ala Ser Cys Trp Cys Ala Tyr
290 295 300
Asn
305

<210> 48
<211> 810
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(810)

<400> 48
atg tcg atc ccc aaa att atg acg gtg tcc aga gac aac gag ggt acg 48
Met Ser Ile Pro Lys Ile Met Thr Val Ser Arg Asp Asn Glu Gly Thr
1 5 10 15

gtg tgt gaa gtc gcg gtg gac aac gga cga cac aga gcg atg att tat 96
Val Cys Glu Val Ala Val Asp Asn Gly Arg His Arg Ala Met Ile Tyr
20 25 30

```
tac cct aag acc acc aac tta gca aac gag cgc gcg gac gtt gtt aag    144
Tyr Pro Lys Thr Thr Asn Leu Ala Asn Glu Arg Ala Asp Val Val Lys
        35                  40                  45

gaa gct ttt gat acc gaa acc cca gtg gac att gta aag caa att gtt    192
Glu Ala Phe Asp Thr Glu Thr Pro Val Asp Ile Val Lys Gln Ile Val
    50                  55                  60

aac gag ggc cta gct ata tcc aaa aaa aat tgc gtc cgt ttg gcg ttg    240
Asn Glu Gly Leu Ala Ile Ser Lys Lys Asn Cys Val Arg Leu Ala Leu
65                  70                  75                  80

tat tta tat ttt tat ttg cag tac gtg tgc ttt gct ctg ctc ctc act    288
Tyr Leu Tyr Phe Tyr Leu Gln Tyr Val Cys Phe Ala Leu Leu Leu Thr
                85                  90                  95

tgg cag tta aac ccg tac atg gac cca ccg ggt ctg gtg ttt gcg gtt    336
Trp Gln Leu Asn Pro Tyr Met Asp Pro Pro Gly Leu Val Phe Ala Val
            100                 105                 110

aac ccc atg ggt cca aaa cat gtc acg aaa cta ccg cac ccg gct att    384
Asn Pro Met Gly Pro Lys His Val Thr Lys Leu Pro His Pro Ala Ile
        115                 120                 125

gtt gcg gta ggt tgt ggg gca gac gcc atc tgt aag aac tgt agc gtc    432
Val Ala Val Gly Cys Gly Ala Asp Ala Ile Cys Lys Asn Cys Ser Val
    130                 135                 140

ccc gat atc aaa acg gag ctt gga atg gtt tac cac aac ggg tct agc    480
Pro Asp Ile Lys Thr Glu Leu Gly Met Val Tyr His Asn Gly Ser Ser
145                 150                 155                 160

gat tct ggt cag cgc gca cac tat ggg ctg gcc ctg tta aag gcg gcc    528
Asp Ser Gly Gln Arg Ala His Tyr Gly Leu Ala Leu Leu Lys Ala Ala
                165                 170                 175

tgg ctt gtc atg gga aat gtg tgt ccg gaa cca gta gtg cgg caa ggc    576
Trp Leu Val Met Gly Asn Val Cys Pro Glu Pro Val Val Arg Gln Gly
            180                 185                 190

gct gca tta ctt ggt cca tgg aac cgg acg gag tgg tcg gat ttt aaa    624
Ala Ala Leu Leu Gly Pro Trp Asn Arg Thr Glu Trp Ser Asp Phe Lys
        195                 200                 205

tcg gca atg gcg gca acc acg ttt tgc gga tcc aga ggc gtt ctg tgg    672
Ser Ala Met Ala Ala Thr Thr Phe Cys Gly Ser Arg Gly Val Leu Trp
    210                 215                 220

tca ccg att cat gaa aaa aac ctc tgt cgc ccc acc tgg aat gat gta    720
Ser Pro Ile His Glu Lys Asn Leu Cys Arg Pro Thr Trp Asn Asp Val
225                 230                 235                 240

att aac aca tca gtt ttt aca aat gaa tca ctc tgt cca aat ata cct    768
Ile Asn Thr Ser Val Phe Thr Asn Glu Ser Leu Cys Pro Asn Ile Pro
                245                 250                 255

gtg gtg ccc gaa agt gta ata gtg ctt aat ggt gat gca tga            810
Val Val Pro Glu Ser Val Ile Val Leu Asn Gly Asp Ala
            260                 265                 270
```

<210> 49
<211> 269
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 49
Met Ser Ile Pro Lys Ile Met Thr Val Ser Arg Asp Asn Glu Gly Thr
1   5   10   15
Val Cys Glu Val Ala Val Asp Asn Gly Arg His Arg Ala Met Ile Tyr
            20   25   30
Tyr Pro Lys Thr Thr Asn Leu Ala Asn Glu Arg Ala Asp Val Val Lys
        35   40   45
Glu Ala Phe Asp Thr Glu Thr Pro Val Asp Ile Val Lys Gln Ile Val
    50   55   60
Asn Glu Gly Leu Ala Ile Ser Lys Lys Asn Cys Val Arg Leu Ala Leu
65   70   75   80
Tyr Leu Tyr Phe Tyr Leu Gln Tyr Val Cys Phe Ala Leu Leu Leu Thr
                85   90   95
Trp Gln Leu Asn Pro Tyr Met Asp Pro Pro Gly Leu Val Phe Ala Val
            100   105   110
Asn Pro Met Gly Pro Lys His Val Thr Lys Leu Pro His Pro Ala Ile
        115   120   125
Val Ala Val Gly Cys Gly Ala Asp Ala Ile Cys Lys Asn Cys Ser Val
    130   135   140
Pro Asp Ile Lys Thr Glu Leu Gly Met Val Tyr His Asn Gly Ser Ser
145   150   155   160
Asp Ser Gly Gln Arg Ala His Tyr Gly Leu Ala Leu Leu Lys Ala Ala
                165   170   175
Trp Leu Val Met Gly Asn Val Cys Pro Glu Pro Val Val Arg Gln Gly
            180   185   190
Ala Ala Leu Leu Gly Pro Trp Asn Arg Thr Glu Trp Ser Asp Phe Lys
        195   200   205
Ser Ala Met Ala Ala Thr Thr Phe Cys Gly Ser Arg Gly Val Leu Trp
    210   215   220
Ser Pro Ile His Glu Lys Asn Leu Cys Arg Pro Thr Trp Asn Asp Val
225   230   235   240
Ile Asn Thr Ser Val Phe Thr Asn Glu Ser Leu Cys Pro Asn Ile Pro
                245   250   255
Val Val Pro Glu Ser Val Ile Val Leu Asn Gly Asp Ala
            260   265

<210> 50
<211> 276
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(276)

<400> 50
atg act gca cac acg aat ggg gtt tta acc acg acg ggc ttt tca aca 48
Met Thr Ala His Thr Asn Gly Val Leu Thr Thr Thr Gly Phe Ser Thr
1   5   10   15

agt cag ccg gaa tcg gtt caa gtt tct cca ttt tat cgc gta att aca 96
Ser Gln Pro Glu Ser Val Gln Val Ser Pro Phe Tyr Arg Val Ile Thr
            20   25   30

aaa cct ccc gtt atg ggc ttg ttt ttt tgc gtg gct atg tgc gtt atc 144
Lys Pro Pro Val Met Gly Leu Phe Phe Cys Val Ala Met Cys Val Ile
35 40 45

gcg ttg gta tgg tac gtg atg cgg agg gtg tgt tgt aag ggg cgc gtt 192
Ala Leu Val Trp Tyr Val Met Arg Arg Val Cys Cys Lys Gly Arg Val
50 55 60

gtt gcc gat tcg tgt cgc gac ccg cgt caa ccc gcg tat gag atg ttg 240
Val Ala Asp Ser Cys Arg Asp Pro Arg Gln Pro Ala Tyr Glu Met Leu
65 70 75 80

aat gtt agg ttg cgt ccc cac gga acc aat cca tag 276
Asn Val Arg Leu Arg Pro His Gly Thr Asn Pro
85 90

<210> 51
<211> 91
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 51
Met Thr Ala His Thr Asn Gly Val Leu Thr Thr Thr Gly Phe Ser Thr
1 5 10 15
Ser Gln Pro Glu Ser Val Gln Val Ser Pro Phe Tyr Arg Val Ile Thr
20 25 30
Lys Pro Pro Val Met Gly Leu Phe Phe Cys Val Ala Met Cys Val Ile
35 40 45
Ala Leu Val Trp Tyr Val Met Arg Arg Val Cys Cys Lys Gly Arg Val
50 55 60
Val Ala Asp Ser Cys Arg Asp Pro Arg Gln Pro Ala Tyr Glu Met Leu
65 70 75 80
Asn Val Arg Leu Arg Pro His Gly Thr Asn Pro
85 90

<210> 52
<211> 1047
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1047)

<400> 52
atg ctt caa aag gac gca aaa ctg att ttt ata tcc tca agt aac tca 48
Met Leu Gln Lys Asp Ala Lys Leu Ile Phe Ile Ser Ser Ser Asn Ser
1 5 10 15

tct gac aaa tca act agt ttt tta ctt aat ttg aag gac gcc cac gaa 96
Ser Asp Lys Ser Thr Ser Phe Leu Leu Asn Leu Lys Asp Ala His Glu
20 25 30

aag atg ctt aac gtg gtg aac tac gta tgt ccg gat cat aaa gat gat 144
Lys Met Leu Asn Val Val Asn Tyr Val Cys Pro Asp His Lys Asp Asp
35 40 45

ttt aac ttg caa gac act gtc gtg gcg tgc ccg tgc tac cgc ctg cat 192

Phe Asn Leu Gln Asp Thr Val Val Ala Cys Pro Cys Tyr Arg Leu His
50 55 60

att cct gcc tat att acg att gat gag acg gta cgg agc acg act aat 240
Ile Pro Ala Tyr Ile Thr Ile Asp Glu Thr Val Arg Ser Thr Thr Asn
65 70 75 80

ctt ttt ttg gaa ggc gca ttc tcc acg gag cta atg ggg gac gct gcc 288
Leu Phe Leu Glu Gly Ala Phe Ser Thr Glu Leu Met Gly Asp Ala Ala
85 90 95

acg tcc gct caa agc atg cat aaa atc gtt agc gac tct tcg ctg tca 336
Thr Ser Ala Gln Ser Met His Lys Ile Val Ser Asp Ser Ser Leu Ser
100 105 110

caa ttg gat ctg tgc cgc gta aag agc acg tca caa gat ata cag ggc 384
Gln Leu Asp Leu Cys Arg Val Lys Ser Thr Ser Gln Asp Ile Gln Gly
115 120 125

gcg atg aag ccg tgt ctc cac gtt tac ata gac ccg gcg tac acc aac 432
Ala Met Lys Pro Cys Leu His Val Tyr Ile Asp Pro Ala Tyr Thr Asn
130 135 140

aac aca gac gcg tcc ggc acc ggc atc ggt gcg gtg att gcg gta aat 480
Asn Thr Asp Ala Ser Gly Thr Gly Ile Gly Ala Val Ile Ala Val Asn
145 150 155 160

cac aag gtg att aaa tgc att tta tta ggc gtg gaa cat ttt ttt cta 528
His Lys Val Ile Lys Cys Ile Leu Leu Gly Val Glu His Phe Phe Leu
165 170 175

aga gat cta acc ggc acc gcc gcg tac cag ata gcg tca tgt gcc gcc 576
Arg Asp Leu Thr Gly Thr Ala Ala Tyr Gln Ile Ala Ser Cys Ala Ala
180 185 190

gcg tta att cga gcg atc gtt acc ctt cac ccg cag atc acg cac gtt 624
Ala Leu Ile Arg Ala Ile Val Thr Leu His Pro Gln Ile Thr His Val
195 200 205

aac gtc gcc gtg gaa ggc aac agc agt caa gat gcc gga gtg gcc ata 672
Asn Val Ala Val Glu Gly Asn Ser Ser Gln Asp Ala Gly Val Ala Ile
210 215 220

gca acc gtg tta aac gag att tgc tcg gtc cct ctt agt ttt tta cac 720
Ala Thr Val Leu Asn Glu Ile Cys Ser Val Pro Leu Ser Phe Leu His
225 230 235 240

cac gtg gac aag aac act ctt ata cgt tcg ccc att tac atg ttg ggg 768
His Val Asp Lys Asn Thr Leu Ile Arg Ser Pro Ile Tyr Met Leu Gly
245 250 255

cca gag aaa gcc aag gcg ttt gaa tct ttt att tac gca ttg aac tcg 816
Pro Glu Lys Ala Lys Ala Phe Glu Ser Phe Ile Tyr Ala Leu Asn Ser
260 265 270

gga acg ttt agc gct agc caa act gtg gtg tct cac act att aag cta 864
Gly Thr Phe Ser Ala Ser Gln Thr Val Val Ser His Thr Ile Lys Leu
275 280 285

tcg ttt gat ccc gta gcg tat cta ata gat caa atc aag gca ata cgt 912
Ser Phe Asp Pro Val Ala Tyr Leu Ile Asp Gln Ile Lys Ala Ile Arg

```
290                 295                 300

tgc att cca cta aaa gac gga ggt cac acg tac tgc gcg aaa caa aaa      960
Cys Ile Pro Leu Lys Asp Gly Gly His Thr Tyr Cys Ala Lys Gln Lys
305                 310                 315                 320

acc atg tcg gac gac gtg ctt gtc gcc gcc gtc atg gcc cac tac atg      1008
Thr Met Ser Asp Asp Val Leu Val Ala Ala Val Met Ala His Tyr Met
            325                 330                 335

gca acc aac gat aaa ttt gtt ttt aaa tcg cta gaa taa                  1047
Ala Thr Asn Asp Lys Phe Val Phe Lys Ser Leu Glu
            340                 345


<210> 53
<211> 348
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 53
Met Leu Gln Lys Asp Ala Lys Leu Ile Phe Ile Ser Ser Ser Asn Ser
1               5                   10                  15
Ser Asp Lys Ser Thr Ser Phe Leu Leu Asn Leu Lys Asp Ala His Glu
            20                  25                  30
Lys Met Leu Asn Val Val Asn Tyr Val Cys Pro Asp His Lys Asp Asp
        35                  40                  45
Phe Asn Leu Gln Asp Thr Val Val Ala Cys Pro Cys Tyr Arg Leu His
    50                  55                  60
Ile Pro Ala Tyr Ile Thr Ile Asp Glu Thr Val Arg Ser Thr Thr Asn
65                  70                  75                  80
Leu Phe Leu Glu Gly Ala Phe Ser Thr Glu Leu Met Gly Asp Ala Ala
                85                  90                  95
Thr Ser Ala Gln Ser Met His Lys Ile Val Ser Asp Ser Ser Leu Ser
            100                 105                 110
Gln Leu Asp Leu Cys Arg Val Lys Ser Thr Ser Gln Asp Ile Gln Gly
        115                 120                 125
Ala Met Lys Pro Cys Leu His Val Tyr Ile Asp Pro Ala Tyr Thr Asn
    130                 135                 140
Asn Thr Asp Ala Ser Gly Thr Gly Ile Gly Ala Val Ile Ala Val Asn
145                 150                 155                 160
His Lys Val Ile Lys Cys Ile Leu Leu Gly Val Glu His Phe Phe Leu
                165                 170                 175
Arg Asp Leu Thr Gly Thr Ala Ala Tyr Gln Ile Ala Ser Cys Ala Ala
            180                 185                 190
Ala Leu Ile Arg Ala Ile Val Thr Leu His Pro Gln Ile Thr His Val
        195                 200                 205
Asn Val Ala Val Glu Gly Asn Ser Ser Gln Asp Ala Gly Val Ala Ile
    210                 215                 220
Ala Thr Val Leu Asn Glu Ile Cys Ser Val Pro Leu Ser Phe Leu His
225                 230                 235                 240
His Val Asp Lys Asn Thr Leu Ile Arg Ser Pro Ile Tyr Met Leu Gly
                245                 250                 255
Pro Glu Lys Ala Lys Ala Phe Glu Ser Phe Ile Tyr Ala Leu Asn Ser
            260                 265                 270
Gly Thr Phe Ser Ala Ser Gln Thr Val Val Ser His Thr Ile Lys Leu
        275                 280                 285
Ser Phe Asp Pro Val Ala Tyr Leu Ile Asp Gln Ile Lys Ala Ile Arg
    290                 295                 300
Cys Ile Pro Leu Lys Asp Gly Gly His Thr Tyr Cys Ala Lys Gln Lys
305                 310                 315                 320
```

Thr Met Ser Asp Asp Val Leu Val Ala Ala Val Met Ala His Tyr Met
325 330 335
Ala Thr Asn Asp Lys Phe Val Phe Lys Ser Leu Glu
340 345

<210> 54
<211> 231
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(231)

<400> 54
atg gag aac gat acg cct aag gac aaa atc tcg gaa gct gac ttt caa 48
Met Glu Asn Asp Thr Pro Lys Asp Lys Ile Ser Glu Ala Asp Phe Gln
1 5 10 15

cag tgt cag gcg ttc ttt cac cgt ccc att aga gat cta att tca tct 96
Gln Cys Gln Ala Phe Phe His Arg Pro Ile Arg Asp Leu Ile Ser Ser
20 25 30

gga gct gac gct tta aac cac ttt agc cta tct gaa tca gac gga cat 144
Gly Ala Asp Ala Leu Asn His Phe Ser Leu Ser Glu Ser Asp Gly His
35 40 45

aaa ttg gaa cgg att gtt ctt ctg ctt gac ctg gtg ggg aca gaa tgt 192
Lys Leu Glu Arg Ile Val Leu Leu Leu Asp Leu Val Gly Thr Glu Cys
50 55 60

ctc tct tat acc acg atc gct gca aag aat gtc aaa tga 231
Leu Ser Tyr Thr Thr Ile Ala Ala Lys Asn Val Lys
65 70 75

<210> 55
<211> 76
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 55
Met Glu Asn Asp Thr Pro Lys Asp Lys Ile Ser Glu Ala Asp Phe Gln
1 5 10 15
Gln Cys Gln Ala Phe Phe His Arg Pro Ile Arg Asp Leu Ile Ser Ser
20 25 30
Gly Ala Asp Ala Leu Asn His Phe Ser Leu Ser Glu Ser Asp Gly His
35 40 45
Lys Leu Glu Arg Ile Val Leu Leu Leu Asp Leu Val Gly Thr Glu Cys
50 55 60
Leu Ser Tyr Thr Thr Ile Ala Ala Lys Asn Val Lys
65 70 75

<210> 56
<211> 654
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(654)

<400> 56
atg tct ctc tta tac cac gat cgc tgc aaa gaa tgt caa atg acg cgc 48
Met Ser Leu Leu Tyr His Asp Arg Cys Lys Glu Cys Gln Met Thr Arg
1 5 10 15

gtc aac agc cca ata tgt cga ttt cat aac gtc tct aac tta tac cag 96
Val Asn Ser Pro Ile Cys Arg Phe His Asn Val Ser Asn Leu Tyr Gln
20 25 30

tgt ttg gat tgt aag cgc tat cac gta tgc gac ggg gga cgc aac tgc 144
Cys Leu Asp Cys Lys Arg Tyr His Val Cys Asp Gly Gly Arg Asn Cys
35 40 45

gtg atc gtg tac act cgc gaa aat cta gtg tgt gat tta acg gga aac 192
Val Ile Val Tyr Thr Arg Glu Asn Leu Val Cys Asp Leu Thr Gly Asn
50 55 60

tgc gtt ttg gat aat gtg cag gac gta tgt tcg tac ggt cct cca gaa 240
Cys Val Leu Asp Asn Val Gln Asp Val Cys Ser Tyr Gly Pro Pro Glu
65 70 75 80

cgc cgc gta ccc gac gcc ttc atc gat ccg ctc gtg tca cac ggc acg 288
Arg Arg Val Pro Asp Ala Phe Ile Asp Pro Leu Val Ser His Gly Thr
85 90 95

agg gaa tgt ctt aaa agc gat ata ctg agg tac ttt gag acg gtc ggt 336
Arg Glu Cys Leu Lys Ser Asp Ile Leu Arg Tyr Phe Glu Thr Val Gly
100 105 110

gtg aaa tct gag gca tat tct acc gtt gtc aag aat gga caa ttg aat 384
Val Lys Ser Glu Ala Tyr Ser Thr Val Val Lys Asn Gly Gln Leu Asn
115 120 125

ggc atc ata ggt aga tta ata gac gct acg ttt aac gag tgc ctt ccg 432
Gly Ile Ile Gly Arg Leu Ile Asp Ala Thr Phe Asn Glu Cys Leu Pro
130 135 140

gta atg agc gac ggc gaa ggt ggc aga gac ctc gcg gcg agc att tac 480
Val Met Ser Asp Gly Glu Gly Gly Arg Asp Leu Ala Ala Ser Ile Tyr
145 150 155 160

atc cac ata att atc tcc ata tac tcc act aaa acg gta tat gat aat 528
Ile His Ile Ile Ile Ser Ile Tyr Ser Thr Lys Thr Val Tyr Asp Asn
165 170 175

ctt cta ttt aaa tgt acg aga aat aaa aaa tac gac cac att gta aaa 576
Leu Leu Phe Lys Cys Thr Arg Asn Lys Lys Tyr Asp His Ile Val Lys
180 185 190

act atc aga gcg caa tgg atg cgc atg gtc tca acc ggc gat ccg tcg 624
Thr Ile Arg Ala Gln Trp Met Arg Met Val Ser Thr Gly Asp Pro Ser
195 200 205

cgg gtc agt gcg acg ggt tgt ttc acg tga 654
Arg Val Ser Ala Thr Gly Cys Phe Thr
210 215

<210> 57
<211> 217
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 57
Met Ser Leu Leu Tyr His Asp Arg Cys Lys Glu Cys Gln Met Thr Arg
1   5   10   15
Val Asn Ser Pro Ile Cys Arg Phe His Asn Val Ser Asn Leu Tyr Gln
            20   25   30
Cys Leu Asp Cys Lys Arg Tyr His Val Cys Asp Gly Gly Arg Asn Cys
        35   40   45
Val Ile Val Tyr Thr Arg Glu Asn Leu Val Cys Asp Leu Thr Gly Asn
    50   55   60
Cys Val Leu Asp Asn Val Gln Asp Val Cys Ser Tyr Gly Pro Pro Glu
65   70   75   80
Arg Arg Val Pro Asp Ala Phe Ile Asp Pro Leu Val Ser His Gly Thr
                85   90   95
Arg Glu Cys Leu Lys Ser Asp Ile Leu Arg Tyr Phe Glu Thr Val Gly
            100   105   110
Val Lys Ser Glu Ala Tyr Ser Thr Val Val Lys Asn Gly Gln Leu Asn
        115   120   125
Gly Ile Ile Gly Arg Leu Ile Asp Ala Thr Phe Asn Glu Cys Leu Pro
    130   135   140
Val Met Ser Asp Gly Glu Gly Gly Arg Asp Leu Ala Ala Ser Ile Tyr
145   150   155   160
Ile His Ile Ile Ile Ser Ile Tyr Ser Thr Lys Thr Val Tyr Asp Asn
                165   170   175
Leu Leu Phe Lys Cys Thr Arg Asn Lys Lys Tyr Asp His Ile Val Lys
            180   185   190
Thr Ile Arg Ala Gln Trp Met Arg Met Val Ser Thr Gly Asp Pro Ser
        195   200   205
Arg Val Ser Ala Thr Gly Cys Phe Thr
    210   215

<210> 58
<211> 1395
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1395)

<400> 58
atg gat gcg cat ggt ctc aac cgg cga tcc gtc gcg ggt cag tgc gac   48
Met Asp Ala His Gly Leu Asn Arg Arg Ser Val Ala Gly Gln Cys Asp
1   5   10   15

ggg ttg ttt cac gtg ata ctt ccg cga ggg ttt atc ctc gcg aac aat   96
Gly Leu Phe His Val Ile Leu Pro Arg Gly Phe Ile Leu Ala Asn Asn
            20   25   30

att acg tgc ggt gaa cgg cag cgg ttt ttt gca cac act tgg ttc gct   144
Ile Thr Cys Gly Glu Arg Gln Arg Phe Phe Ala His Thr Trp Phe Ala
        35   40   45

gca tct gga cgc acg tct aag act tta tac gtg tgg gga cgg gta ttt 192
Ala Ser Gly Arg Thr Ser Lys Thr Leu Tyr Val Trp Gly Arg Val Phe
50 55 60

caa aac acc gac ccg ggc cgc ggg gac ggt ccg tcc ggg ccg tgg tcc 240
Gln Asn Thr Asp Pro Gly Arg Gly Asp Gly Pro Ser Gly Pro Trp Ser
65 70 75 80

gga ctg gcg att agt ctg cct ctg ttt acc aca aat gga aaa ttt cat 288
Gly Leu Ala Ile Ser Leu Pro Leu Phe Thr Thr Asn Gly Lys Phe His
85 90 95

ccg ttt gat gta gtt ata ctc aag gcc gat acg cct gac tct gga agc 336
Pro Phe Asp Val Val Ile Leu Lys Ala Asp Thr Pro Asp Ser Gly Ser
100 105 110

tcg tgg acc gtg aag ttc ttg tat atg tca tta att gcg gct tac aga 384
Ser Trp Thr Val Lys Phe Leu Tyr Met Ser Leu Ile Ala Ala Tyr Arg
115 120 125

aac gca atg cga ggt tta aaa gat aaa gtt tcg caa tgt acc gat gcc 432
Asn Ala Met Arg Gly Leu Lys Asp Lys Val Ser Gln Cys Thr Asp Ala
130 135 140

gcc gtt gac ggt gag gtt cat cct cta acc gtc tta aaa gaa gcg ttg 480
Ala Val Asp Gly Glu Val His Pro Leu Thr Val Leu Lys Glu Ala Leu
145 150 155 160

gta tca ccg gac act gct acg cga ccc gtg tcc gcg tgc aac cct cta 528
Val Ser Pro Asp Thr Ala Thr Arg Pro Val Ser Ala Cys Asn Pro Leu
165 170 175

cag atg ttg acc gga ctc tta cag tct agg gta cgg gac gac tac gtg 576
Gln Met Leu Thr Gly Leu Leu Gln Ser Arg Val Arg Asp Asp Tyr Val
180 185 190

aca cac cac cgt gcg ctc gaa cgc cca ggt aat gtg agg gga caa gta 624
Thr His His Arg Ala Leu Glu Arg Pro Gly Asn Val Arg Gly Gln Val
195 200 205

atc gcc ccg acg cgc acc gag atg cca aac gga tcg cca agt cgt gta 672
Ile Ala Pro Thr Arg Thr Glu Met Pro Asn Gly Ser Pro Ser Arg Val
210 215 220

agg ctt gga ttc cgc cct ccc aaa caa gcc aac tat cca aag acg tgg 720
Arg Leu Gly Phe Arg Pro Pro Lys Gln Ala Asn Tyr Pro Lys Thr Trp
225 230 235 240

gcg cag gcg cgt cac gtt ttc tcg tct cgc gca tat tac gtg tgc gta 768
Ala Gln Ala Arg His Val Phe Ser Ser Arg Ala Tyr Tyr Val Cys Val
245 250 255

tat gat aac gaa gaa cta gat acc aag tgg cag cgg caa gat ccg cgg 816
Tyr Asp Asn Glu Glu Leu Asp Thr Lys Trp Gln Arg Gln Asp Pro Arg
260 265 270

ccg ttg ccg cta gat tgg tcc gat ccg gtc gcg tac ctg tta gag ggc 864
Pro Leu Pro Leu Asp Trp Ser Asp Pro Val Ala Tyr Leu Leu Glu Gly
275 280 285

gat ttg ttt tta gga gcc aag cag aat gcg ttt gtc gat tct ata gaa 912

Asp Leu Phe Leu Gly Ala Lys Gln Asn Ala Phe Val Asp Ser Ile Glu
290 295 300

aag acg tgc agg tgt cag aac tat acc att aag caa ttt ttt ccg gtt 960
Lys Thr Cys Arg Cys Gln Asn Tyr Thr Ile Lys Gln Phe Phe Pro Val
305 310 315 320

ttg ata aat agg gac aac gaa aca gtc gac tta att aag gag cat ttt 1008
Leu Ile Asn Arg Asp Asn Glu Thr Val Asp Leu Ile Lys Glu His Phe
325 330 335

ata gag gcg tgc ttc gtg att aga aac cag gtg tca gag agg agc gct 1056
Ile Glu Ala Cys Phe Val Ile Arg Asn Gln Val Ser Glu Arg Ser Ala
340 345 350

tgg gta aag gcg gcg ctg ttt cgc aac gat agt aac acg tat tgg aag 1104
Trp Val Lys Ala Ala Leu Phe Arg Asn Asp Ser Asn Thr Tyr Trp Lys
355 360 365

gat gtt ttg gga tta tgg gag cat ggg cct cat aag ctg ggt acg gct 1152
Asp Val Leu Gly Leu Trp Glu His Gly Pro His Lys Leu Gly Thr Ala
370 375 380

ata aaa cta cca aca tcg gaa cct tgc aat gcc gac gta aac tgg agc 1200
Ile Lys Leu Pro Thr Ser Glu Pro Cys Asn Ala Asp Val Asn Trp Ser
385 390 395 400

tgg ctc ctg tgt gac gag gat ata act cgg tca att agc ggg cag tct 1248
Trp Leu Leu Cys Asp Glu Asp Ile Thr Arg Ser Ile Ser Gly Gln Ser
405 410 415

act gtt tgc tta gtt gtc tcg cct acc ctg acc gcc tgg ctg gtg ctc 1296
Thr Val Cys Leu Val Val Ser Pro Thr Leu Thr Ala Trp Leu Val Leu
420 425 430

ccg ggg ggc ttt gtt att aaa ggc cgc tac gac cta tca agc gag gat 1344
Pro Gly Gly Phe Val Ile Lys Gly Arg Tyr Asp Leu Ser Ser Glu Asp
435 440 445

tta atg ttt gtg gct tcg aga tat ggc cac cca gcg tcg tca cat tct 1392
Leu Met Phe Val Ala Ser Arg Tyr Gly His Pro Ala Ser Ser His Ser
450 455 460

taa 1395

465

<210> 59
<211> 464
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 59
Met Asp Ala His Gly Leu Asn Arg Arg Ser Val Ala Gly Gln Cys Asp
1 5 10 15
Gly Leu Phe His Val Ile Leu Pro Arg Gly Phe Ile Leu Ala Asn Asn
20 25 30
Ile Thr Cys Gly Glu Arg Gln Arg Phe Phe Ala His Thr Trp Phe Ala
35 40 45
Ala Ser Gly Arg Thr Ser Lys Thr Leu Tyr Val Trp Gly Arg Val Phe

```
     50                  55                  60
Gln Asn Thr Asp Pro Gly Arg Gly Asp Gly Pro Ser Gly Pro Trp Ser
 65                  70                  75                  80
Gly Leu Ala Ile Ser Leu Pro Leu Phe Thr Thr Asn Gly Lys Phe His
                 85                  90                  95
Pro Phe Asp Val Val Ile Leu Lys Ala Asp Thr Pro Asp Ser Gly Ser
            100                 105                 110
Ser Trp Thr Val Lys Phe Leu Tyr Met Ser Leu Ile Ala Ala Tyr Arg
        115                 120                 125
Asn Ala Met Arg Gly Leu Lys Asp Lys Val Ser Gln Cys Thr Asp Ala
    130                 135                 140
Ala Val Asp Gly Glu Val His Pro Leu Thr Val Leu Lys Glu Ala Leu
145                 150                 155                 160
Val Ser Pro Asp Thr Ala Thr Arg Pro Val Ser Ala Cys Asn Pro Leu
                165                 170                 175
Gln Met Leu Thr Gly Leu Leu Gln Ser Arg Val Arg Asp Asp Tyr Val
            180                 185                 190
Thr His His Arg Ala Leu Glu Arg Pro Gly Asn Val Arg Gly Gln Val
        195                 200                 205
Ile Ala Pro Thr Arg Thr Glu Met Pro Asn Gly Ser Pro Ser Arg Val
    210                 215                 220
Arg Leu Gly Phe Arg Pro Pro Lys Gln Ala Asn Tyr Pro Lys Thr Trp
225                 230                 235                 240
Ala Gln Ala Arg His Val Phe Ser Ser Arg Ala Tyr Tyr Val Cys Val
                245                 250                 255
Tyr Asp Asn Glu Glu Leu Asp Thr Lys Trp Gln Arg Gln Asp Pro Arg
            260                 265                 270
Pro Leu Pro Leu Asp Trp Ser Asp Pro Val Ala Tyr Leu Leu Glu Gly
        275                 280                 285
Asp Leu Phe Leu Gly Ala Lys Gln Asn Ala Phe Val Asp Ser Ile Glu
    290                 295                 300
Lys Thr Cys Arg Cys Gln Asn Tyr Thr Ile Lys Gln Phe Phe Pro Val
305                 310                 315                 320
Leu Ile Asn Arg Asp Asn Glu Thr Val Asp Leu Ile Lys Glu His Phe
                325                 330                 335
Ile Glu Ala Cys Phe Val Ile Arg Asn Gln Val Ser Glu Arg Ser Ala
            340                 345                 350
Trp Val Lys Ala Ala Leu Phe Arg Asn Asp Ser Asn Thr Tyr Trp Lys
        355                 360                 365
Asp Val Leu Gly Leu Trp Glu His Gly Pro His Lys Leu Gly Thr Ala
    370                 375                 380
Ile Lys Leu Pro Thr Ser Glu Pro Cys Asn Ala Asp Val Asn Trp Ser
385                 390                 395                 400
Trp Leu Leu Cys Asp Glu Asp Ile Thr Arg Ser Ile Ser Gly Gln Ser
                405                 410                 415
Thr Val Cys Leu Val Val Ser Pro Thr Leu Thr Ala Trp Leu Val Leu
            420                 425                 430
Pro Gly Gly Phe Val Ile Lys Gly Arg Tyr Asp Leu Ser Ser Glu Asp
        435                 440                 445
Leu Met Phe Val Ala Ser Arg Tyr Gly His Pro Ala Ser Ser His Ser
    450                 455                 460


<210> 60
<211> 1011
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
```

<222> (1)..(1011)

<400> 60

```
atg gcc acc cag cgt cgt cac att ctt aaa tcg ttt tta aac aag gaa      48
Met Ala Thr Gln Arg Arg His Ile Leu Lys Ser Phe Leu Asn Lys Glu
1               5                   10                  15

tgc ata tgg ttg cga cac ccg ggt acg tcc gcg ttt gtt cgg gta tac      96
Cys Ile Trp Leu Arg His Pro Gly Thr Ser Ala Phe Val Arg Val Tyr
            20                  25                  30

acc gcg acc act gcg cat tct gcc gtt ttt gac ccg ccg gta act agc     144
Thr Ala Thr Thr Ala His Ser Ala Val Phe Asp Pro Pro Val Thr Ser
        35                  40                  45

gaa aat gcg atg tca ctt aac ttt tta aat gtt atg atc gta atc atg     192
Glu Asn Ala Met Ser Leu Asn Phe Leu Asn Val Met Ile Val Ile Met
    50                  55                  60

aaa cca aaa gaa ttt ggc ccg tgc gtt acc gtg tac atg aat gga gat     240
Lys Pro Lys Glu Phe Gly Pro Cys Val Thr Val Tyr Met Asn Gly Asp
65                  70                  75                  80

att cta gat ttt tgt gcc acg gaa tct gtc gcc ata agg gac gtg cct     288
Ile Leu Asp Phe Cys Ala Thr Glu Ser Val Ala Ile Arg Asp Val Pro
                85                  90                  95

ggt agg gcg gac ctg tgt tta att cgt ttt ggt acc ctt tct aat gcg     336
Gly Arg Ala Asp Leu Cys Leu Ile Arg Phe Gly Thr Leu Ser Asn Ala
            100                 105                 110

ccg agg agc gtt ccg ata ccc ggg ccg ttg aac cca cat ccg cga gaa     384
Pro Arg Ser Val Pro Ile Pro Gly Pro Leu Asn Pro His Pro Arg Glu
        115                 120                 125

acc gtg ccc ggg cta aca aaa cag gaa att ata tac act tcg caa aca     432
Thr Val Pro Gly Leu Thr Lys Gln Glu Ile Ile Tyr Thr Ser Gln Thr
    130                 135                 140

gtg cca aga gga cag ata cca gat gcc ata aag ggg aaa gag ttc cac     480
Val Pro Arg Gly Gln Ile Pro Asp Ala Ile Lys Gly Lys Glu Phe His
145                 150                 155                 160

caa ata aat ccg ttt ttg tgg ttt gac gga ggg gcg ttt tgg caa ctg     528
Gln Ile Asn Pro Phe Leu Trp Phe Asp Gly Gly Ala Phe Trp Gln Leu
                165                 170                 175

ttc ctc tct gtg gat ttt atg ctg ctc tgt ccc gca ctc gac aca gtt     576
Phe Leu Ser Val Asp Phe Met Leu Leu Cys Pro Ala Leu Asp Thr Val
            180                 185                 190

ccg tcc ctg gcc aga atc gtt ggg ctt ctt aca cag tgc gat aag agc     624
Pro Ser Leu Ala Arg Ile Val Gly Leu Leu Thr Gln Cys Asp Lys Ser
        195                 200                 205

acg tgt aaa att tgt acg ggg gcc cac gta cac gtt aac ccg tat cgc     672
Thr Cys Lys Ile Cys Thr Gly Ala His Val His Val Asn Pro Tyr Arg
    210                 215                 220

gga tac acg cca cct gac tcg caa ggg acc tca ccc tcg tgc ccc tgc     720
Gly Tyr Thr Pro Pro Asp Ser Gln Gly Thr Ser Pro Ser Cys Pro Cys
```

225 230 235 240

ctt atc tcg tgc ggg gcc agg cgc gcg gcg gat gtc ctg gtt acc gga 768
Leu Ile Ser Cys Gly Ala Arg Arg Ala Ala Asp Val Leu Val Thr Gly
245 250 255

cac gtt aat ctt ttg ggc ctg ctc ttt gac ccc aaa gcc tcc ccc aaa 816
His Val Asn Leu Leu Gly Leu Leu Phe Asp Pro Lys Ala Ser Pro Lys
260 265 270

gtg acc aag ctg cgt tta aaa aga aac cca cgc ccg gta ccg ata gag 864
Val Thr Lys Leu Arg Leu Lys Arg Asn Pro Arg Pro Val Pro Ile Glu
275 280 285

gac gcc atg tcg ggc gtc acg gcc gaa ggg acc gag gtg caa ccc act 912
Asp Ala Met Ser Gly Val Thr Ala Glu Gly Thr Glu Val Gln Pro Thr
290 295 300

tcg cta ccg tgg gcc ctc att cgc ctg ccg gat tta gcc agt cgc gtg 960
Ser Leu Pro Trp Ala Leu Ile Arg Leu Pro Asp Leu Ala Ser Arg Val
305 310 315 320

atg cta tac ggc tgc cag aac tta aaa agc atc tgc tta cgt tct tat 1008
Met Leu Tyr Gly Cys Gln Asn Leu Lys Ser Ile Cys Leu Arg Ser Tyr
325 330 335

tga 1011

<210> 61
<211> 336
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 61
Met Ala Thr Gln Arg Arg His Ile Leu Lys Ser Phe Leu Asn Lys Glu
1 5 10 15
Cys Ile Trp Leu Arg His Pro Gly Thr Ser Ala Phe Val Arg Val Tyr
20 25 30
Thr Ala Thr Thr Ala His Ser Ala Val Phe Asp Pro Pro Val Thr Ser
35 40 45
Glu Asn Ala Met Ser Leu Asn Phe Leu Asn Val Met Ile Val Ile Met
50 55 60
Lys Pro Lys Glu Phe Gly Pro Cys Val Thr Val Tyr Met Asn Gly Asp
65 70 75 80
Ile Leu Asp Phe Cys Ala Thr Glu Ser Val Ala Ile Arg Asp Val Pro
85 90 95
Gly Arg Ala Asp Leu Cys Leu Ile Arg Phe Gly Thr Leu Ser Asn Ala
100 105 110
Pro Arg Ser Val Pro Ile Pro Gly Pro Leu Asn Pro His Pro Arg Glu
115 120 125
Thr Val Pro Gly Leu Thr Lys Gln Glu Ile Ile Tyr Thr Ser Gln Thr
130 135 140
Val Pro Arg Gly Gln Ile Pro Asp Ala Ile Lys Gly Lys Glu Phe His
145 150 155 160
Gln Ile Asn Pro Phe Leu Trp Phe Asp Gly Gly Ala Phe Trp Gln Leu
165 170 175
Phe Leu Ser Val Asp Phe Met Leu Leu Cys Pro Ala Leu Asp Thr Val
180 185 190
Pro Ser Leu Ala Arg Ile Val Gly Leu Leu Thr Gln Cys Asp Lys Ser

195 200 205
Thr Cys Lys Ile Cys Thr Gly Ala His Val His Val Asn Pro Tyr Arg
210 215 220
Gly Tyr Thr Pro Pro Asp Ser Gln Gly Thr Ser Pro Ser Cys Pro Cys
225 230 235 240
Leu Ile Ser Cys Gly Ala Arg Arg Ala Ala Asp Val Leu Val Thr Gly
245 250 255
His Val Asn Leu Leu Gly Leu Leu Phe Asp Pro Lys Ala Ser Pro Lys
260 265 270
Val Thr Lys Leu Arg Leu Lys Arg Asn Pro Arg Pro Val Pro Ile Glu
275 280 285
Asp Ala Met Ser Gly Val Thr Ala Glu Gly Thr Glu Val Gln Pro Thr
290 295 300
Ser Leu Pro Trp Ala Leu Ile Arg Leu Pro Asp Leu Ala Ser Arg Val
305 310 315 320
Met Leu Tyr Gly Cys Gln Asn Leu Lys Ser Ile Cys Leu Arg Ser Tyr
325 330 335

<210> 62
<211> 984
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(984)

<400> 62
atg ttg tta acc agc tat cgc gaa cgt ctt caa aat aac ttg cgc gtg 48
Met Leu Leu Thr Ser Tyr Arg Glu Arg Leu Gln Asn Asn Leu Arg Val
1 5 10 15

gtc acg gac ggt ggt tgc gaa aac tgg ttt cgg caa ccg ccc gtt att 96
Val Thr Asp Gly Gly Cys Glu Asn Trp Phe Arg Gln Pro Pro Val Ile
20 25 30

ata tcg ggc aac gac aag acc gaa cga atg gcc cac cca tgc ttg gga 144
Ile Ser Gly Asn Asp Lys Thr Glu Arg Met Ala His Pro Cys Leu Gly
35 40 45

gtt att cac gcg gtt aat gca tat agt tct gtt tta gac gat tat ctt 192
Val Ile His Ala Val Asn Ala Tyr Ser Ser Val Leu Asp Asp Tyr Leu
50 55 60

caa acg tac cgc aga gtt caa gaa ccc atg ccg gcc cct acg ttg gga 240
Gln Thr Tyr Arg Arg Val Gln Glu Pro Met Pro Ala Pro Thr Leu Gly
65 70 75 80

aag ccc cga att tct agc cac gct acg ttg ccc cgg tta acc gag gaa 288
Lys Pro Arg Ile Ser Ser His Ala Thr Leu Pro Arg Leu Thr Glu Glu
85 90 95

ctc aca aac tac ctt aaa caa aca tgt tgt cgg gtc caa atg gca aac 336
Leu Thr Asn Tyr Leu Lys Gln Thr Cys Cys Arg Val Gln Met Ala Asn
100 105 110

gcc aag gac cag tac atg gaa tac caa tcg gcc caa cgg acc cac gaa 384
Ala Lys Asp Gln Tyr Met Glu Tyr Gln Ser Ala Gln Arg Thr His Glu
115 120 125

gct ttc cta gag tgc ccc gtt tat gca gaa ctg cgg cag ttt tta gcc 432
Ala Phe Leu Glu Cys Pro Val Tyr Ala Glu Leu Arg Gln Phe Leu Ala
130 135 140

aac ctg tcg tca ttt tta aat ggg agt tac gtg ccc ggg gtt tgt tgc 480
Asn Leu Ser Ser Phe Leu Asn Gly Ser Tyr Val Pro Gly Val Cys Cys
145 150 155 160

ctt gag ccc ttt caa caa caa tta atc atg cac acg ttt tat ttt atc 528
Leu Glu Pro Phe Gln Gln Gln Leu Ile Met His Thr Phe Tyr Phe Ile
165 170 175

gcg tct atc aaa gca ccc gaa aag aca cac cag tta ttt gcc acg ttt 576
Ala Ser Ile Lys Ala Pro Glu Lys Thr His Gln Leu Phe Ala Thr Phe
180 185 190

aag caa cac ttt ggt tta ttt gaa acc acg gac gac gtg tta cag acg 624
Lys Gln His Phe Gly Leu Phe Glu Thr Thr Asp Asp Val Leu Gln Thr
195 200 205

ttt aag cag aaa gcc agc gtt ttt gtt atc ccg cgt cgt cac ggg aaa 672
Phe Lys Gln Lys Ala Ser Val Phe Val Ile Pro Arg Arg His Gly Lys
210 215 220

aca tgg atc gtg gtc gcc att atc agc gtt ctt ctc tcg tcg gtt gaa 720
Thr Trp Ile Val Val Ala Ile Ile Ser Val Leu Leu Ser Ser Val Glu
225 230 235 240

aac gtt cac gtg ggt tac gtg gcc cac caa aaa cac gtc gcc aat gcc 768
Asn Val His Val Gly Tyr Val Ala His Gln Lys His Val Ala Asn Ala
245 250 255

gtt ttc tcc gag gtc atc gcc acg cta tcc agg tgg ttt ccg gcg aag 816
Val Phe Ser Glu Val Ile Ala Thr Leu Ser Arg Trp Phe Pro Ala Lys
260 265 270

aac ctg aac ata aag aag gaa aac gga acc atc gtg tac gcg agc ccc 864
Asn Leu Asn Ile Lys Lys Glu Asn Gly Thr Ile Val Tyr Ala Ser Pro
275 280 285

gga cgg cgg ccg agt tcg ctg atg tgt gcg acg tgc ttc aat aag aac 912
Gly Arg Arg Pro Ser Ser Leu Met Cys Ala Thr Cys Phe Asn Lys Asn
290 295 300

gta agc aga tgc ttt tta agt tct ggc agc cgt ata gca tca cgc gac 960
Val Ser Arg Cys Phe Leu Ser Ser Gly Ser Arg Ile Ala Ser Arg Asp
305 310 315 320

tgg cta aat ccg gca ggc gaa tga 984
Trp Leu Asn Pro Ala Gly Glu
325

<210> 63
<211> 327
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 63
Met Leu Leu Thr Ser Tyr Arg Glu Arg Leu Gln Asn Asn Leu Arg Val

1 5 10 15
Val Thr Asp Gly Gly Cys Glu Asn Trp Phe Arg Gln Pro Pro Val Ile
20 25 30
Ile Ser Gly Asn Asp Lys Thr Glu Arg Met Ala His Pro Cys Leu Gly
35 40 45
Val Ile His Ala Val Asn Ala Tyr Ser Ser Val Leu Asp Asp Tyr Leu
50 55 60
Gln Thr Tyr Arg Arg Val Gln Glu Pro Met Pro Ala Pro Thr Leu Gly
65 70 75 80
Lys Pro Arg Ile Ser Ser His Ala Thr Leu Pro Arg Leu Thr Glu Glu
85 90 95
Leu Thr Asn Tyr Leu Lys Gln Thr Cys Cys Arg Val Gln Met Ala Asn
100 105 110
Ala Lys Asp Gln Tyr Met Glu Tyr Gln Ser Ala Gln Arg Thr His Glu
115 120 125
Ala Phe Leu Glu Cys Pro Val Tyr Ala Glu Leu Arg Gln Phe Leu Ala
130 135 140
Asn Leu Ser Ser Phe Leu Asn Gly Ser Tyr Val Pro Gly Val Cys Cys
145 150 155 160
Leu Glu Pro Phe Gln Gln Gln Leu Ile Met His Thr Phe Tyr Phe Ile
165 170 175
Ala Ser Ile Lys Ala Pro Glu Lys Thr His Gln Leu Phe Ala Thr Phe
180 185 190
Lys Gln His Phe Gly Leu Phe Glu Thr Thr Asp Asp Val Leu Gln Thr
195 200 205
Phe Lys Gln Lys Ala Ser Val Phe Val Ile Pro Arg Arg His Gly Lys
210 215 220
Thr Trp Ile Val Val Ala Ile Ile Ser Val Leu Leu Ser Ser Val Glu
225 230 235 240
Asn Val His Val Gly Tyr Val Ala His Gln Lys His Val Ala Asn Ala
245 250 255
Val Phe Ser Glu Val Ile Ala Thr Leu Ser Arg Trp Phe Pro Ala Lys
260 265 270
Asn Leu Asn Ile Lys Lys Glu Asn Gly Thr Ile Val Tyr Ala Ser Pro
275 280 285
Gly Arg Arg Pro Ser Ser Leu Met Cys Ala Thr Cys Phe Asn Lys Asn
290 295 300
Val Ser Arg Cys Phe Leu Ser Ser Gly Ser Arg Ile Ala Ser Arg Asp
305 310 315 320
Trp Leu Asn Pro Ala Gly Glu
325

<210> 64
<211> 984
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(984)

<400> 64
atg ttc cct tca agt ttt ttg aat aac ggg cac ccc gaa acg gaa cgc 48
Met Phe Pro Ser Ser Phe Leu Asn Asn Gly His Pro Glu Thr Glu Arg
1 5 10 15

cgc ttc gtt aaa ggc gtt cag tta gct tta gac ctg tgt gac aac act 96
Arg Phe Val Lys Gly Val Gln Leu Ala Leu Asp Leu Cys Asp Asn Thr
20 25 30

```
ccg gga cag ttt aaa cta gtt gaa aca cct ctt aat agt ttt ctt ctg    144
Pro Gly Gln Phe Lys Leu Val Glu Thr Pro Leu Asn Ser Phe Leu Leu
        35                  40                  45

gta tcc aac gtt ctg ccg gaa tcg cgc ccg gtt aga gac tgt ccg cag    192
Val Ser Asn Val Leu Pro Glu Ser Arg Pro Val Arg Asp Cys Pro Gln
    50                  55                  60

ccg gaa ggg ttt gac ttt gaa cac att cac ctc cca aaa cta aca cgc    240
Pro Glu Gly Phe Asp Phe Glu His Ile His Leu Pro Lys Leu Thr Arg
65                  70                  75                  80

atg cag cgt gtc ctg ggg cga tac tgc gac cat gtt aac aac gac gac    288
Met Gln Arg Val Leu Gly Arg Tyr Cys Asp His Val Asn Asn Asp Asp
                85                  90                  95

acg tgc gtt aac gta aag gca agt tcc tcg aat tca cag ggt gcc ttg    336
Thr Cys Val Asn Val Lys Ala Ser Ser Ser Asn Ser Gln Gly Ala Leu
            100                 105                 110

ttt tat ctg ccg tat gga cag gac gaa tgg aat tgg gcg ctc acg tta    384
Phe Tyr Leu Pro Tyr Gly Gln Asp Glu Trp Asn Trp Ala Leu Thr Leu
        115                 120                 125

agg aaa gac aag ttg gtt aaa atg gct gta gag ggc ttg tca aat ccc    432
Arg Lys Asp Lys Leu Val Lys Met Ala Val Glu Gly Leu Ser Asn Pro
    130                 135                 140

acg acc tgg aaa ggt tta gag ccc gtg gat cct tta ccg ctc ata tgg    480
Thr Thr Trp Lys Gly Leu Glu Pro Val Asp Pro Leu Pro Leu Ile Trp
145                 150                 155                 160

ctt ctg ttt tac ggt tcc cgg tcg ttc tgt cgg gaa cca gag tgc cta    528
Leu Leu Phe Tyr Gly Ser Arg Ser Phe Cys Arg Glu Pro Glu Cys Leu
                165                 170                 175

tat gaa cgc aat ttt ggt atg aag gga ccc ata ctc tta ccg cca cat    576
Tyr Glu Arg Asn Phe Gly Met Lys Gly Pro Ile Leu Leu Pro Pro His
            180                 185                 190

atg tat gcc ccc caa aag gac gta atg act ttt gtc cat cat gta att    624
Met Tyr Ala Pro Gln Lys Asp Val Met Thr Phe Val His His Val Ile
        195                 200                 205

aag tac gtt aaa ttt tta tac gtg aac gcc ggt ggg ggt ctt gaa act    672
Lys Tyr Val Lys Phe Leu Tyr Val Asn Ala Gly Gly Gly Leu Glu Thr
    210                 215                 220

gaa ccg tcc ccg ccg ttc gag gcc tcg cgg ttg cgc gca gcc atc gct    720
Glu Pro Ser Pro Pro Phe Glu Ala Ser Arg Leu Arg Ala Ala Ile Ala
225                 230                 235                 240

cgt ctc ggg gac gtg gaa gcg gat gac gca tac ctg tcc gca aag tgc    768
Arg Leu Gly Asp Val Glu Ala Asp Asp Ala Tyr Leu Ser Ala Lys Cys
                245                 250                 255

atg ttg tgt cac ctg tac aag caa aac gat acg att tcg att cat gaa    816
Met Leu Cys His Leu Tyr Lys Gln Asn Asp Thr Ile Ser Ile His Glu
            260                 265                 270
```

```
aca cac gtg ggc gga gtc atc gca tta ggc gga gac ggt gcg aga tat    864
Thr His Val Gly Gly Val Ile Ala Leu Gly Gly Asp Gly Ala Arg Tyr
        275                 280                 285

ata acg tct agt gtt cgg gct caa cgg tgc acg agt cgg gga gat ttt    912
Ile Thr Ser Ser Val Arg Ala Gln Arg Cys Thr Ser Arg Gly Asp Phe
    290                 295                 300

gtt tta atc ccc ctg tac aac att gaa ggg ctc gta agc atg ata agg    960
Val Leu Ile Pro Leu Tyr Asn Ile Glu Gly Leu Val Ser Met Ile Arg
305                 310                 315                 320

gaa cat ggc ctc ggc agc agc taa                                    984
Glu His Gly Leu Gly Ser Ser
                325
```

```
<210> 65
<211> 327
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 65
Met Phe Pro Ser Ser Phe Leu Asn Asn Gly His Pro Glu Thr Glu Arg
1               5                   10                  15
Arg Phe Val Lys Gly Val Gln Leu Ala Leu Asp Leu Cys Asp Asn Thr
            20                  25                  30
Pro Gly Gln Phe Lys Leu Val Glu Thr Pro Leu Asn Ser Phe Leu Leu
        35                  40                  45
Val Ser Asn Val Leu Pro Glu Ser Arg Pro Val Arg Asp Cys Pro Gln
    50                  55                  60
Pro Glu Gly Phe Asp Phe Glu His Ile His Leu Pro Lys Leu Thr Arg
65                  70                  75                  80
Met Gln Arg Val Leu Gly Arg Tyr Cys Asp His Val Asn Asn Asp Asp
                85                  90                  95
Thr Cys Val Asn Val Lys Ala Ser Ser Ser Asn Ser Gln Gly Ala Leu
            100                 105                 110
Phe Tyr Leu Pro Tyr Gly Gln Asp Glu Trp Asn Trp Ala Leu Thr Leu
        115                 120                 125
Arg Lys Asp Lys Leu Val Lys Met Ala Val Glu Gly Leu Ser Asn Pro
    130                 135                 140
Thr Thr Trp Lys Gly Leu Glu Pro Val Asp Pro Leu Pro Leu Ile Trp
145                 150                 155                 160
Leu Leu Phe Tyr Gly Ser Arg Ser Phe Cys Arg Glu Pro Glu Cys Leu
                165                 170                 175
Tyr Glu Arg Asn Phe Gly Met Lys Gly Pro Ile Leu Leu Pro Pro His
            180                 185                 190
Met Tyr Ala Pro Gln Lys Asp Val Met Thr Phe Val His His Val Ile
        195                 200                 205
Lys Tyr Val Lys Phe Leu Tyr Val Asn Ala Gly Gly Gly Leu Glu Thr
    210                 215                 220
Glu Pro Ser Pro Pro Phe Glu Ala Ser Arg Leu Arg Ala Ala Ile Ala
225                 230                 235                 240
Arg Leu Gly Asp Val Glu Ala Asp Asp Ala Tyr Leu Ser Ala Lys Cys
                245                 250                 255
Met Leu Cys His Leu Tyr Lys Gln Asn Asp Thr Ile Ser Ile His Glu
            260                 265                 270
Thr His Val Gly Gly Val Ile Ala Leu Gly Gly Asp Gly Ala Arg Tyr
        275                 280                 285
Ile Thr Ser Ser Val Arg Ala Gln Arg Cys Thr Ser Arg Gly Asp Phe
    290                 295                 300
```

Val Leu Ile Pro Leu Tyr Asn Ile Glu Gly Leu Val Ser Met Ile Arg
305 310 315 320
Glu His Gly Leu Gly Ser Ser
325

<210> 66
<211> 450
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(450)

<400> 66
atg gcc tcg gca gca gct aaa aaa atg tta att aag tcg gag ctc gag 48
Met Ala Ser Ala Ala Ala Lys Lys Met Leu Ile Lys Ser Glu Leu Glu
1 5 10 15

tcg gaa atc aac aaa aaa ctg tcc atc tcc gta ttt gac agg ttt ggg 96
Ser Glu Ile Asn Lys Lys Leu Ser Ile Ser Val Phe Asp Arg Phe Gly
20 25 30

gcc gac agt gcc gtg ttt aac gcg cag tat aag gga acc agg gaa tcg 144
Ala Asp Ser Ala Val Phe Asn Ala Gln Tyr Lys Gly Thr Arg Glu Ser
35 40 45

ctg cgg tcg tac aac agc cta aaa aag aag gac gat ctg gcg acc gtt 192
Leu Arg Ser Tyr Asn Ser Leu Lys Lys Lys Asp Asp Leu Ala Thr Val
50 55 60

gtc gga acg cta gaa acg tcg ctg cgt gaa aaa caa agc gaa ttg gga 240
Val Gly Thr Leu Glu Thr Ser Leu Arg Glu Lys Gln Ser Glu Leu Gly
65 70 75 80

tta cta aag ggg ttt aac agg aaa aaa att gaa gag ttt gac gct gtg 288
Leu Leu Lys Gly Phe Asn Arg Lys Lys Ile Glu Glu Phe Asp Ala Val
85 90 95

gcg gac gcg gtt cgc gac ctc aag gac gag ctg tac gga gaa ctg gag 336
Ala Asp Ala Val Arg Asp Leu Lys Asp Glu Leu Tyr Gly Glu Leu Glu
100 105 110

att cta ggt acg ctt gac aat gaa tct gtt ccc gtg gaa gaa gag tcc 384
Ile Leu Gly Thr Leu Asp Asn Glu Ser Val Pro Val Glu Glu Glu Ser
115 120 125

ccc aag gac gac att att agg tgg aaa ttg gag cgt ctg ccc aga gtg 432
Pro Lys Asp Asp Ile Ile Arg Trp Lys Leu Glu Arg Leu Pro Arg Val
130 135 140

tgc ccc aaa agc cct tga 450
Cys Pro Lys Ser Pro
145 150

<210> 67
<211> 149
<212> PRT

<213> Macaca mulatta rhadinovirus 17577

<400> 67
Met Ala Ser Ala Ala Ala Lys Lys Met Leu Ile Lys Ser Glu Leu Glu
1 5 10 15
Ser Glu Ile Asn Lys Lys Leu Ser Ile Ser Val Phe Asp Arg Phe Gly
20 25 30
Ala Asp Ser Ala Val Phe Asn Ala Gln Tyr Lys Gly Thr Arg Glu Ser
35 40 45
Leu Arg Ser Tyr Asn Ser Leu Lys Lys Lys Asp Asp Leu Ala Thr Val
50 55 60
Val Gly Thr Leu Glu Thr Ser Leu Arg Glu Lys Gln Ser Glu Leu Gly
65 70 75 80
Leu Leu Lys Gly Phe Asn Arg Lys Lys Ile Glu Glu Phe Asp Ala Val
85 90 95
Ala Asp Ala Val Arg Asp Leu Lys Asp Glu Leu Tyr Gly Glu Leu Glu
100 105 110
Ile Leu Gly Thr Leu Asp Asn Glu Ser Val Pro Val Glu Glu Glu Ser
115 120 125
Pro Lys Asp Asp Ile Ile Arg Trp Lys Leu Glu Arg Leu Pro Arg Val
130 135 140
Cys Pro Lys Ser Pro
145

<210> 68
<211> 1308
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1308)

<400> 68
atg aat ctg ttc ccg tgg aag aag agt ccc cca agg acg aca tta tta 48
Met Asn Leu Phe Pro Trp Lys Lys Ser Pro Pro Arg Thr Thr Leu Leu
1 5 10 15

ggt gga aat tgg agc gtc tgc cca gag tgt gcc cca aaa gcc ctt gat 96
Gly Gly Asn Trp Ser Val Cys Pro Glu Cys Ala Pro Lys Ala Leu Asp
20 25 30

ccc att ccc aag gtt cag act gac gtc gac aga aca gca tcg tcc cat 144
Pro Ile Pro Lys Val Gln Thr Asp Val Asp Arg Thr Ala Ser Ser His
35 40 45

ata acc gtc att aaa aca cgt aag acg atc gcc caa ctg aag ata cct 192
Ile Thr Val Ile Lys Thr Arg Lys Thr Ile Ala Gln Leu Lys Ile Pro
50 55 60

aac aac tgg ggc cag tgt agt cac cag gcg acg gac tgg acc gcc gtg 240
Asn Asn Trp Gly Gln Cys Ser His Gln Ala Thr Asp Trp Thr Ala Val
65 70 75 80

ctc gga cgc ggc tcg tat ggt gtg gtg agg tcc atg tct ctc ggc cgc 288
Leu Gly Arg Gly Ser Tyr Gly Val Val Arg Ser Met Ser Leu Gly Arg
85 90 95

tgc gtt aag cat ttt ggc agc cgg cgt gag ttt ttt tac gag tgc att 336

```
Cys Val Lys His Phe Gly Ser Arg Arg Glu Phe Phe Tyr Glu Cys Ile
            100                 105                 110

ttt aac gat ata gta cgc gcc tgc cgg gag aaa cat ccc ctg aac cgc     384
Phe Asn Asp Ile Val Arg Ala Cys Arg Glu Lys His Pro Leu Asn Arg
        115                 120                 125

ggg ggt gac cgt ata cta tgt ttc cta gag ccg tgc gta cca tgt cgc     432
Gly Gly Asp Arg Ile Leu Cys Phe Leu Glu Pro Cys Val Pro Cys Arg
    130                 135                 140

gcc ctg ata ttc ccg cag tta aca ggg aat ctg cta aac gcg gat ctt     480
Ala Leu Ile Phe Pro Gln Leu Thr Gly Asn Leu Leu Asn Ala Asp Leu
145                 150                 155                 160

aaa cac gtg aac cct gaa cgg ctg gcc gtt gaa ttc tct gag ctg agg     528
Lys His Val Asn Pro Glu Arg Leu Ala Val Glu Phe Ser Glu Leu Arg
                165                 170                 175

gaa ggc gtt agt ttt cta aat aat ata tgt ggc atc gtt cac tgt gac     576
Glu Gly Val Ser Phe Leu Asn Asn Ile Cys Gly Ile Val His Cys Asp
            180                 185                 190

atc agt cca gaa aat ata ctg ata aag ggg gaa ctg aca act gcg tac     624
Ile Ser Pro Glu Asn Ile Leu Ile Lys Gly Glu Leu Thr Thr Ala Tyr
        195                 200                 205

ggg aga ctt atg atc gga gat cta ggg tcc gcc tct tta cac acg gga     672
Gly Arg Leu Met Ile Gly Asp Leu Gly Ser Ala Ser Leu His Thr Gly
    210                 215                 220

acc cct tgg acc gga gtg atg gtg acc tcc aaa ctc ggg ttc gtg cag     720
Thr Pro Trp Thr Gly Val Met Val Thr Ser Lys Leu Gly Phe Val Gln
225                 230                 235                 240

cac acg tac cat ttt aag gca ccg gcc aga ttt atc tgt aag cac att     768
His Thr Tyr His Phe Lys Ala Pro Ala Arg Phe Ile Cys Lys His Ile
                245                 250                 255

tac cgg ccg tcg tgt ctc ctc tac cgg tgt ctg ctg tcg tgc gcc ggg     816
Tyr Arg Pro Ser Cys Leu Leu Tyr Arg Cys Leu Leu Ser Cys Ala Gly
            260                 265                 270

ggc ccg cag gcg cat atg cta aat cag ccg ttc caa atc act cca caa     864
Gly Pro Gln Ala His Met Leu Asn Gln Pro Phe Gln Ile Thr Pro Gln
        275                 280                 285

ctc ggt ctc aca att gac ata tcg tcc ctg ggt tat agt ttg cta gca     912
Leu Gly Leu Thr Ile Asp Ile Ser Ser Leu Gly Tyr Ser Leu Leu Ala
    290                 295                 300

tgc cta gag aaa tat ctt cag cca gct gac cca ttt ccc cag cag gga     960
Cys Leu Glu Lys Tyr Leu Gln Pro Ala Asp Pro Phe Pro Gln Gln Gly
305                 310                 315                 320

gcg ttg gcg gac gct tcc tcc gaa tcc gcc cac cca ttg ttc tat ttg     1008
Ala Leu Ala Asp Ala Ser Ser Glu Ser Ala His Pro Leu Phe Tyr Leu
                325                 330                 335

cgt tgc atg gtg cca aga gta gtc atc gcc gag att ttt tct gtt gcc     1056
Arg Cys Met Val Pro Arg Val Val Ile Ala Glu Ile Phe Ser Val Ala
```

340 345 350

tgg gac gtt cca ctc gat tta ggc att gac tca tct ggc cac gcg cca 1104
Trp Asp Val Pro Leu Asp Leu Gly Ile Asp Ser Ser Gly His Ala Pro
355 360 365

gct att ccc ctg aga gaa gcg tac agg cgg ttt ttt gcc aac cag tgt 1152
Ala Ile Pro Leu Arg Glu Ala Tyr Arg Arg Phe Phe Ala Asn Gln Cys
370 375 380

agt tta tat agg gcg caa tac aaa gag gat gcg tta gaa aac gca tcc 1200
Ser Leu Tyr Arg Ala Gln Tyr Lys Glu Asp Ala Leu Glu Asn Ala Ser
385 390 395 400

tcg cgg ctg tgt aac tca aaa ctt aaa cta gtt ctc cag aag ctg ttg 1248
Ser Arg Leu Cys Asn Ser Lys Leu Lys Leu Val Leu Gln Lys Leu Leu
405 410 415

gtc agg gac tac ttt agt cat tgc gga aac tgc gga gat cat gga ttt 1296
Val Arg Asp Tyr Phe Ser His Cys Gly Asn Cys Gly Asp His Gly Phe
420 425 430

ttt ctc aga tga 1308
Phe Leu Arg
435

<210> 69
<211> 435
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 69
Met Asn Leu Phe Pro Trp Lys Lys Ser Pro Pro Arg Thr Thr Leu Leu
1 5 10 15
Gly Gly Asn Trp Ser Val Cys Pro Glu Cys Ala Pro Lys Ala Leu Asp
20 25 30
Pro Ile Pro Lys Val Gln Thr Asp Val Asp Arg Thr Ala Ser Ser His
35 40 45
Ile Thr Val Ile Lys Thr Arg Lys Thr Ile Ala Gln Leu Lys Ile Pro
50 55 60
Asn Asn Trp Gly Gln Cys Ser His Gln Ala Thr Asp Trp Thr Ala Val
65 70 75 80
Leu Gly Arg Gly Ser Tyr Gly Val Val Arg Ser Met Ser Leu Gly Arg
85 90 95
Cys Val Lys His Phe Gly Ser Arg Arg Glu Phe Phe Tyr Glu Cys Ile
100 105 110
Phe Asn Asp Ile Val Arg Ala Cys Arg Glu Lys His Pro Leu Asn Arg
115 120 125
Gly Gly Asp Arg Ile Leu Cys Phe Leu Glu Pro Cys Val Pro Cys Arg
130 135 140
Ala Leu Ile Phe Pro Gln Leu Thr Gly Asn Leu Leu Asn Ala Asp Leu
145 150 155 160
Lys His Val Asn Pro Glu Arg Leu Ala Val Glu Phe Ser Glu Leu Arg
165 170 175
Glu Gly Val Ser Phe Leu Asn Asn Ile Cys Gly Ile Val His Cys Asp
180 185 190
Ile Ser Pro Glu Asn Ile Leu Ile Lys Gly Glu Leu Thr Thr Ala Tyr
195 200 205
Gly Arg Leu Met Ile Gly Asp Leu Gly Ser Ala Ser Leu His Thr Gly
210 215 220

Thr Pro Trp Thr Gly Val Met Val Thr Ser Lys Leu Gly Phe Val Gln
225 230 235 240
His Thr Tyr His Phe Lys Ala Pro Ala Arg Phe Ile Cys Lys His Ile
245 250 255
Tyr Arg Pro Ser Cys Leu Leu Tyr Arg Cys Leu Leu Ser Cys Ala Gly
260 265 270
Gly Pro Gln Ala His Met Leu Asn Gln Pro Phe Gln Ile Thr Pro Gln
275 280 285
Leu Gly Leu Thr Ile Asp Ile Ser Ser Leu Gly Tyr Ser Leu Leu Ala
290 295 300
Cys Leu Glu Lys Tyr Leu Gln Pro Ala Asp Pro Phe Pro Gln Gln Gly
305 310 315 320
Ala Leu Ala Asp Ala Ser Ser Glu Ser Ala His Pro Leu Phe Tyr Leu
325 330 335
Arg Cys Met Val Pro Arg Val Val Ile Ala Glu Ile Phe Ser Val Ala
340 345 350
Trp Asp Val Pro Leu Asp Leu Gly Ile Asp Ser Ser Gly His Ala Pro
355 360 365
Ala Ile Pro Leu Arg Glu Ala Tyr Arg Arg Phe Phe Ala Asn Gln Cys
370 375 380
Ser Leu Tyr Arg Ala Gln Tyr Lys Glu Asp Ala Leu Glu Asn Ala Ser
385 390 395 400
Ser Arg Leu Cys Asn Ser Lys Leu Lys Leu Val Leu Gln Lys Leu Leu
405 410 415
Val Arg Asp Tyr Phe Ser His Cys Gly Asn Cys Gly Asp His Gly Phe
420 425 430
Phe Leu Arg
435

<210> 70
<211> 1443
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1443)

<400> 70
atg gat ttt ttc tca gat gag ccg atg gtt cag gag atg gcg ctt ctc 48
Met Asp Phe Phe Ser Asp Glu Pro Met Val Gln Glu Met Ala Leu Leu
1 5 10 15

gac atc gat gag cag cag cgg ctc ctc tcg aaa atg agc ctg gcc aac 96
Asp Ile Asp Glu Gln Gln Arg Leu Leu Ser Lys Met Ser Leu Ala Asn
20 25 30

ttt tta aaa cat gag cga gta agg gcg ttt ttt agc gat aac aaa aag 144
Phe Leu Lys His Glu Arg Val Arg Ala Phe Phe Ser Asp Asn Lys Lys
35 40 45

gaa ata agc atg ccg gct ata cgg ttc gtg tat aat ttt tat cta ttc 192
Glu Ile Ser Met Pro Ala Ile Arg Phe Val Tyr Asn Phe Tyr Leu Phe
50 55 60

gcc aag gtg gga gat ttt atc ggc aac acc gac gtg tac gat ttt tac 240
Ala Lys Val Gly Asp Phe Ile Gly Asn Thr Asp Val Tyr Asp Phe Tyr
65 70 75 80

gtc act tgc gtg ttc agg ggg agg cgc ctg acg cgc ctg tcg gaa gtg 288
Val Thr Cys Val Phe Arg Gly Arg Arg Leu Thr Arg Leu Ser Glu Val
85 90 95

tac gac gcg tgc cta aac atg cac ccg cac gat cga cac cac gtg tgt 336
Tyr Asp Ala Cys Leu Asn Met His Pro His Asp Arg His His Val Cys
100 105 110

gca ttg ata gaa cag gtc acg cgc ggc caa aac atc aat cct ctg tgg 384
Ala Leu Ile Glu Gln Val Thr Arg Gly Gln Asn Ile Asn Pro Leu Trp
115 120 125

gac gct ctg agg gac ggc ata att tcg tcg tca aaa ttt cac tgg gcc 432
Asp Ala Leu Arg Asp Gly Ile Ile Ser Ser Ser Lys Phe His Trp Ala
130 135 140

ata aaa caa cag aat tcg tcc aaa aaa att ttt aac ccg tgg cct ata 480
Ile Lys Gln Gln Asn Ser Ser Lys Lys Ile Phe Asn Pro Trp Pro Ile
145 150 155 160

gtc aac aat cac ttt gta gcg ggc ccg ctc gcg ttt gga ctg cgt tgc 528
Val Asn Asn His Phe Val Ala Gly Pro Leu Ala Phe Gly Leu Arg Cys
165 170 175

gag gaa gtg gtt aaa aag ata ctg gcg acg ttg ctg cat cca ggc gag 576
Glu Glu Val Val Lys Lys Ile Leu Ala Thr Leu Leu His Pro Gly Glu
180 185 190

gcg cac tgt gaa aac tac gga ttc atg cag agt cct ctc aac ggg gtt 624
Ala His Cys Glu Asn Tyr Gly Phe Met Gln Ser Pro Leu Asn Gly Val
195 200 205

ttt ggc gtc tcc ttg gat ttt gga att aac gtc agg tct gac cca aaa 672
Phe Gly Val Ser Leu Asp Phe Gly Ile Asn Val Arg Ser Asp Pro Lys
210 215 220

gac ggt ttg gag ttt cac cca gac tgc aaa atc tat gaa ata aaa tgc 720
Asp Gly Leu Glu Phe His Pro Asp Cys Lys Ile Tyr Glu Ile Lys Cys
225 230 235 240

cgg ttt aag tac act ttt tcc aag atg gag tgt gac ccg att tac gct 768
Arg Phe Lys Tyr Thr Phe Ser Lys Met Glu Cys Asp Pro Ile Tyr Ala
245 250 255

gcg tat gct aaa ctt tat cag aag ccc agc atg cag acg ctt aag ggg 816
Ala Tyr Ala Lys Leu Tyr Gln Lys Pro Ser Met Gln Thr Leu Lys Gly
260 265 270

ttt ttg tac tcc ata tct aaa ccg gcg atc gag ttt gtc gga gag gac 864
Phe Leu Tyr Ser Ile Ser Lys Pro Ala Ile Glu Phe Val Gly Glu Asp
275 280 285

agg ctc ccc agt gaa tcg gac tat ctt gtg gca tat gac aaa gaa tgg 912
Arg Leu Pro Ser Glu Ser Asp Tyr Leu Val Ala Tyr Asp Lys Glu Trp
290 295 300

gag gtg tgt ccg cgg aaa aag aga cgc tta act gca gta cac cat cta 960
Glu Val Cys Pro Arg Lys Lys Arg Arg Leu Thr Ala Val His His Leu
305 310 315 320

gtt aaa aag tgc atg att cac aac tct acg gcg cct tct gat gtg tat 1008

Val Lys Lys Cys Met Ile His Asn Ser Thr Ala Pro Ser Asp Val Tyr
325 330 335

ata ttg tca gat ccg cag gaa acc gga ggc caa att aat att aaa gct 1056
Ile Leu Ser Asp Pro Gln Glu Thr Gly Gly Gln Ile Asn Ile Lys Ala
340 345 350

cat ctg agc gcc aac cta ttt ata aac gtc agg cat ccg tat tat tat 1104
His Leu Ser Ala Asn Leu Phe Ile Asn Val Arg His Pro Tyr Tyr Tyr
355 360 365

caa gtg ttg ctc cag tct ctc gtt gta cag gag tac atc agt ctc tcc 1152
Gln Val Leu Leu Gln Ser Leu Val Val Gln Glu Tyr Ile Ser Leu Ser
370 375 380

aag gga act aaa aat ttg gga acc cag aaa aac ttt ata gcg act ggc 1200
Lys Gly Thr Lys Asn Leu Gly Thr Gln Lys Asn Phe Ile Ala Thr Gly
385 390 395 400

ttt ttt aga aaa cgg cag ttt caa gac ccg agc tgc tgc acg atc ggt 1248
Phe Phe Arg Lys Arg Gln Phe Gln Asp Pro Ser Cys Cys Thr Ile Gly
405 410 415

gaa ttt gcc ccg ttg gat cca cac gta gag ata ccg acc ctt tta atc 1296
Glu Phe Ala Pro Leu Asp Pro His Val Glu Ile Pro Thr Leu Leu Ile
420 425 430

gtg aca ccg gtg tat ttt ccc agc gtg gcc aaa cac caa ctg gtg aag 1344
Val Thr Pro Val Tyr Phe Pro Ser Val Ala Lys His Gln Leu Val Lys
435 440 445

cag gcg acc gaa ttc tgg gcg gct agt gct cgt gag gca ttt ccc gag 1392
Gln Ala Thr Glu Phe Trp Ala Ala Ser Ala Arg Glu Ala Phe Pro Glu
450 455 460

cta cca tgg gat tta tcc tct ctg tgt gca aac gcc cca cca aca ccg 1440
Leu Pro Trp Asp Leu Ser Ser Leu Cys Ala Asn Ala Pro Pro Thr Pro
465 470 475 480

tag 1443

<210> 71
<211> 480
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 71
Met Asp Phe Phe Ser Asp Glu Pro Met Val Gln Glu Met Ala Leu Leu
1 5 10 15
Asp Ile Asp Glu Gln Gln Arg Leu Leu Ser Lys Met Ser Leu Ala Asn
20 25 30
Phe Leu Lys His Glu Arg Val Arg Ala Phe Phe Ser Asp Asn Lys Lys
35 40 45
Glu Ile Ser Met Pro Ala Ile Arg Phe Val Tyr Asn Phe Tyr Leu Phe
50 55 60
Ala Lys Val Gly Asp Phe Ile Gly Asn Thr Asp Val Tyr Asp Phe Tyr
65 70 75 80
Val Thr Cys Val Phe Arg Gly Arg Arg Leu Thr Arg Leu Ser Glu Val
85 90 95

```
Tyr Asp Ala Cys Leu Asn Met His Pro His Asp Arg His His Val Cys
            100                 105                 110
Ala Leu Ile Glu Gln Val Thr Arg Gly Gln Asn Ile Asn Pro Leu Trp
        115                 120                 125
Asp Ala Leu Arg Asp Gly Ile Ile Ser Ser Ser Lys Phe His Trp Ala
    130                 135                 140
Ile Lys Gln Gln Asn Ser Ser Lys Lys Ile Phe Asn Pro Trp Pro Ile
145                 150                 155                 160
Val Asn Asn His Phe Val Ala Gly Pro Leu Ala Phe Gly Leu Arg Cys
                165                 170                 175
Glu Glu Val Val Lys Lys Ile Leu Ala Thr Leu Leu His Pro Gly Glu
            180                 185                 190
Ala His Cys Glu Asn Tyr Gly Phe Met Gln Ser Pro Leu Asn Gly Val
        195                 200                 205
Phe Gly Val Ser Leu Asp Phe Gly Ile Asn Val Arg Ser Asp Pro Lys
    210                 215                 220
Asp Gly Leu Glu Phe His Pro Asp Cys Lys Ile Tyr Glu Ile Lys Cys
225                 230                 235                 240
Arg Phe Lys Tyr Thr Phe Ser Lys Met Glu Cys Asp Pro Ile Tyr Ala
                245                 250                 255
Ala Tyr Ala Lys Leu Tyr Gln Lys Pro Ser Met Gln Thr Leu Lys Gly
            260                 265                 270
Phe Leu Tyr Ser Ile Ser Lys Pro Ala Ile Glu Phe Val Gly Glu Asp
        275                 280                 285
Arg Leu Pro Ser Glu Ser Asp Tyr Leu Val Ala Tyr Asp Lys Glu Trp
    290                 295                 300
Glu Val Cys Pro Arg Lys Lys Arg Arg Leu Thr Ala Val His His Leu
305                 310                 315                 320
Val Lys Lys Cys Met Ile His Asn Ser Thr Ala Pro Ser Asp Val Tyr
                325                 330                 335
Ile Leu Ser Asp Pro Gln Glu Thr Gly Gly Gln Ile Asn Ile Lys Ala
            340                 345                 350
His Leu Ser Ala Asn Leu Phe Ile Asn Val Arg His Pro Tyr Tyr Tyr
        355                 360                 365
Gln Val Leu Leu Gln Ser Leu Val Val Gln Glu Tyr Ile Ser Leu Ser
    370                 375                 380
Lys Gly Thr Lys Asn Leu Gly Thr Gln Lys Asn Phe Ile Ala Thr Gly
385                 390                 395                 400
Phe Phe Arg Lys Arg Gln Phe Gln Asp Pro Ser Cys Cys Thr Ile Gly
                405                 410                 415
Glu Phe Ala Pro Leu Asp Pro His Val Glu Ile Pro Thr Leu Leu Ile
            420                 425                 430
Val Thr Pro Val Tyr Phe Pro Ser Val Ala Lys His Gln Leu Val Lys
        435                 440                 445
Gln Ala Thr Glu Phe Trp Ala Ala Ser Ala Arg Glu Ala Phe Pro Glu
    450                 455                 460
Leu Pro Trp Asp Leu Ser Ser Leu Cys Ala Asn Ala Pro Pro Thr Pro
465                 470                 475                 480
```

<210> 72
<211> 210
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(210)

<400> 72

atg gga ttt atc ctc tct gtg tgc aaa cgc ccc acc aac acc gta gat 48
Met Gly Phe Ile Leu Ser Val Cys Lys Arg Pro Thr Asn Thr Val Asp
1 5 10 15

gtg aag ggg gag ccc ata gat gta tcc aaa gaa ttc gat cct att ata 96
Val Lys Gly Glu Pro Ile Asp Val Ser Lys Glu Phe Asp Pro Ile Ile
20 25 30

gga gaa gaa agc att gtc ttg tta acg gca gat ggg act gcc ccc gcg 144
Gly Glu Glu Ser Ile Val Leu Leu Thr Ala Asp Gly Thr Ala Pro Ala
35 40 45

gcg ctg tac aaa ccc aaa acc aag cca tcc aaa cat aaa aac aat aaa 192
Ala Leu Tyr Lys Pro Lys Thr Lys Pro Ser Lys His Lys Asn Asn Lys
50 55 60

ttg tca gat ttt gtt taa 210
Leu Ser Asp Phe Val
65 70

<210> 73
<211> 69
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 73
Met Gly Phe Ile Leu Ser Val Cys Lys Arg Pro Thr Asn Thr Val Asp
1 5 10 15
Val Lys Gly Glu Pro Ile Asp Val Ser Lys Glu Phe Asp Pro Ile Ile
20 25 30
Gly Glu Glu Ser Ile Val Leu Leu Thr Ala Asp Gly Thr Ala Pro Ala
35 40 45
Ala Leu Tyr Lys Pro Lys Thr Lys Pro Ser Lys His Lys Asn Asn Lys
50 55 60
Leu Ser Asp Phe Val
65

<210> 74
<211> 1137
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1137)

<400> 74
atg aaa att tca cgg agc gac tcg ttt atc tta tcg tca tgg gta aaa 48
Met Lys Ile Ser Arg Ser Asp Ser Phe Ile Leu Ser Ser Trp Val Lys
1 5 10 15

cta ctg gtt att ctt gga ctt atg ttt ata atg tca gcg gta gtg cca 96
Leu Leu Val Ile Leu Gly Leu Met Phe Ile Met Ser Ala Val Val Pro
20 25 30

ctg acc gcc aca ttc ccg gga ctt gga ttt ccg tgc tac ttt aac acg 144
Leu Thr Ala Thr Phe Pro Gly Leu Gly Phe Pro Cys Tyr Phe Asn Thr
35 40 45

ttg gtc aac tac agc gcg tta aac cta acg gtc aga agt tct gct aaa 192
Leu Val Asn Tyr Ser Ala Leu Asn Leu Thr Val Arg Ser Ser Ala Lys
50 55 60

cat ctg acg ccg act ttg ttt ttg gaa gca cca gaa atg ttt gtc tac 240
His Leu Thr Pro Thr Leu Phe Leu Glu Ala Pro Glu Met Phe Val Tyr
65 70 75 80

atc tcg tgg gcg ttt ttg gtg gac gga tat ctg ctg tgt tac tat gcg 288
Ile Ser Trp Ala Phe Leu Val Asp Gly Tyr Leu Leu Cys Tyr Tyr Ala
85 90 95

tgg gcc atc ctg gcc ata ttc aag gcc aag cgc gtg cac gcg aca aca 336
Trp Ala Ile Leu Ala Ile Phe Lys Ala Lys Arg Val His Ala Thr Thr
100 105 110

atg acc agc ctc cag aca tgg atc gtg ctc ata ggc tcc cac agt gta 384
Met Thr Ser Leu Gln Thr Trp Ile Val Leu Ile Gly Ser His Ser Val
115 120 125

gtt ttt atg tct att ttg agg ctg tgg acc atc cag tta ttt att cac 432
Val Phe Met Ser Ile Leu Arg Leu Trp Thr Ile Gln Leu Phe Ile His
130 135 140

gtt ctg tcg tac aaa cat ata ctt tta gcc tcg ttt gta tac tgt ata 480
Val Leu Ser Tyr Lys His Ile Leu Leu Ala Ser Phe Val Tyr Cys Ile
145 150 155 160

cat ttt tgc cta tcg ttt acg cac gtc cag gct atg ata tcg tgc aat 528
His Phe Cys Leu Ser Phe Thr His Val Gln Ala Met Ile Ser Cys Asn
165 170 175

tcg gcc acc tgg tct cta cgc gtc tta gag cag cag att cct gag aat 576
Ser Ala Thr Trp Ser Leu Arg Val Leu Glu Gln Gln Ile Pro Glu Asn
180 185 190

agc ttg cta gac acc ctg ttg cgt tac gga aaa ccg att ggc gcc aat 624
Ser Leu Leu Asp Thr Leu Leu Arg Tyr Gly Lys Pro Ile Gly Ala Asn
195 200 205

ctt tat ctg tcc tta ata gcc atg gag atg tta gta ttc tcc ctc gga 672
Leu Tyr Leu Ser Leu Ile Ala Met Glu Met Leu Val Phe Ser Leu Gly
210 215 220

acg atg atg gct att gga aac agt ttc tat atg ctc gtt tcc gat att 720
Thr Met Met Ala Ile Gly Asn Ser Phe Tyr Met Leu Val Ser Asp Ile
225 230 235 240

gtc ttt ggt tca ata aac ttg ttt ttt gtc ctt aca ata gct tgg tat 768
Val Phe Gly Ser Ile Asn Leu Phe Phe Val Leu Thr Ile Ala Trp Tyr
245 250 255

atc aac aca gaa cta ttt cta gta aag tac cta aag cac cag atc gga 816
Ile Asn Thr Glu Leu Phe Leu Val Lys Tyr Leu Lys His Gln Ile Gly
260 265 270

ttc tac gtt ggc gtt ttt gtc agt tac ctg att ctg ctt ctt ccc gtc 864
Phe Tyr Val Gly Val Phe Val Ser Tyr Leu Ile Leu Leu Leu Pro Val
275 280 285

```
gtt cga tac gac aag gta ttc ata tct gcc agc ctg cac aaa gtc ata     912
Val Arg Tyr Asp Lys Val Phe Ile Ser Ala Ser Leu His Lys Val Ile
    290                 295                 300

gcc gtg aac atc tct atg att ccg atc act tgc atc cta gcc atc atc     960
Ala Val Asn Ile Ser Met Ile Pro Ile Thr Cys Ile Leu Ala Ile Ile
305                 310                 315                 320

tta aga att atc aga aac gat tgg aaa tgg tgc gca aag tcg ccc gaa    1008
Leu Arg Ile Ile Arg Asn Asp Trp Lys Trp Cys Ala Lys Ser Pro Glu
                325                 330                 335

tac gca ccc ctt cct caa ggc cct aaa gag aag acg aca aaa gtt aag    1056
Tyr Ala Pro Leu Pro Gln Gly Pro Lys Glu Lys Thr Thr Lys Val Lys
            340                 345                 350

tac tcc ccg gaa ctg aac gcg tta tat gaa aca gaa gaa gac gtg agc    1104
Tyr Ser Pro Glu Leu Asn Ala Leu Tyr Glu Thr Glu Glu Asp Val Ser
        355                 360                 365

gat tac gag gat gca tat cca aaa tac ata tga                        1137
Asp Tyr Glu Asp Ala Tyr Pro Lys Tyr Ile
    370                 375


<210> 75
<211> 378
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 75
Met Lys Ile Ser Arg Ser Asp Ser Phe Ile Leu Ser Ser Trp Val Lys
1               5                   10                  15
Leu Leu Val Ile Leu Gly Leu Met Phe Ile Met Ser Ala Val Val Pro
            20                  25                  30
Leu Thr Ala Thr Phe Pro Gly Leu Gly Phe Pro Cys Tyr Phe Asn Thr
        35                  40                  45
Leu Val Asn Tyr Ser Ala Leu Asn Leu Thr Val Arg Ser Ser Ala Lys
    50                  55                  60
His Leu Thr Pro Thr Leu Phe Leu Glu Ala Pro Glu Met Phe Val Tyr
65                  70                  75                  80
Ile Ser Trp Ala Phe Leu Val Asp Gly Tyr Leu Leu Cys Tyr Tyr Ala
                85                  90                  95
Trp Ala Ile Leu Ala Ile Phe Lys Ala Lys Arg Val His Ala Thr Thr
            100                 105                 110
Met Thr Ser Leu Gln Thr Trp Ile Val Leu Ile Gly Ser His Ser Val
        115                 120                 125
Val Phe Met Ser Ile Leu Arg Leu Trp Thr Ile Gln Leu Phe Ile His
    130                 135                 140
Val Leu Ser Tyr Lys His Ile Leu Leu Ala Ser Phe Val Tyr Cys Ile
145                 150                 155                 160
His Phe Cys Leu Ser Phe Thr His Val Gln Ala Met Ile Ser Cys Asn
                165                 170                 175
Ser Ala Thr Trp Ser Leu Arg Val Leu Glu Gln Gln Ile Pro Glu Asn
            180                 185                 190
Ser Leu Leu Asp Thr Leu Leu Arg Tyr Gly Lys Pro Ile Gly Ala Asn
        195                 200                 205
Leu Tyr Leu Ser Leu Ile Ala Met Glu Met Leu Val Phe Ser Leu Gly
    210                 215                 220
Thr Met Met Ala Ile Gly Asn Ser Phe Tyr Met Leu Val Ser Asp Ile
225                 230                 235                 240
```

Val Phe Gly Ser Ile Asn Leu Phe Phe Val Leu Thr Ile Ala Trp Tyr
245 250 255
Ile Asn Thr Glu Leu Phe Leu Val Lys Tyr Leu Lys His Gln Ile Gly
260 265 270
Phe Tyr Val Gly Val Phe Val Ser Tyr Leu Ile Leu Leu Leu Pro Val
275 280 285
Val Arg Tyr Asp Lys Val Phe Ile Ser Ala Ser Leu His Lys Val Ile
290 295 300
Ala Val Asn Ile Ser Met Ile Pro Ile Thr Cys Ile Leu Ala Ile Ile
305 310 315 320
Leu Arg Ile Ile Arg Asn Asp Trp Lys Trp Cys Ala Lys Ser Pro Glu
325 330 335
Tyr Ala Pro Leu Pro Gln Gly Pro Lys Glu Lys Thr Thr Lys Val Lys
340 345 350
Tyr Ser Pro Glu Leu Asn Ala Leu Tyr Glu Thr Glu Glu Asp Val Ser
355 360 365
Asp Tyr Glu Asp Ala Tyr Pro Lys Tyr Ile
370 375

<210> 76
<211> 1407
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1407)

<400> 76
atg aac gcc cgg gag gtg gca ctc acg gga cat gtt ttg cac ata tcg 48
Met Asn Ala Arg Glu Val Ala Leu Thr Gly His Val Leu His Ile Ser
1 5 10 15

ctg cat agc acg cac gag cgc gag aaa tta ata atc tgg cag gtt cat 96
Leu His Ser Thr His Glu Arg Glu Lys Leu Ile Ile Trp Gln Val His
20 25 30

tta ctt gta tgt caa caa tgc gga att cag gga gat gcc gca tat cta 144
Leu Leu Val Cys Gln Gln Cys Gly Ile Gln Gly Asp Ala Ala Tyr Leu
35 40 45

ttt gtc acc gaa aca tta agc aat act gac tgg gga aat ata ccg gcg 192
Phe Val Thr Glu Thr Leu Ser Asn Thr Asp Trp Gly Asn Ile Pro Ala
50 55 60

ata aac cgt cac gca ccg tcc ata aat gag cac ggc cgt aat tat atg 240
Ile Asn Arg His Ala Pro Ser Ile Asn Glu His Gly Arg Asn Tyr Met
65 70 75 80

cag tgg gaa ctc cgt act cgt tta cgg aat ccc atc att caa ttg tta 288
Gln Trp Glu Leu Arg Thr Arg Leu Arg Asn Pro Ile Ile Gln Leu Leu
85 90 95

agt cgc cag ccc ggt gcg gtt aac gta agg gtc agc gag ccg aat atg 336
Ser Arg Gln Pro Gly Ala Val Asn Val Arg Val Ser Glu Pro Asn Met
100 105 110

gta ata gtg ggc tgc gaa cga gcg ttg gat cac tcg tgt tcg gtg cgc 384
Val Ile Val Gly Cys Glu Arg Ala Leu Asp His Ser Cys Ser Val Arg

115 120 125

gtg act gga gcc tat ctt cat tgc gat acc act atg gac ttt agt ttg 432
Val Thr Gly Ala Tyr Leu His Cys Asp Thr Thr Met Asp Phe Ser Leu
130 135 140

gat tct gtt gtg tcc cca acc cgc gaa ttt tgg ttc tca gag atg ttt 480
Asp Ser Val Val Ser Pro Thr Arg Glu Phe Trp Phe Ser Glu Met Phe
145 150 155 160

tct cac tgt tta gtt tcc aac att gaa gtt tac ctt aaa aca acg ggc 528
Ser His Cys Leu Val Ser Asn Ile Glu Val Tyr Leu Lys Thr Thr Gly
165 170 175

ggg tta tac tat agg gca tcg agt gcc acg caa tgc cga aaa agg gcg 576
Gly Leu Tyr Tyr Arg Ala Ser Ser Ala Thr Gln Cys Arg Lys Arg Ala
180 185 190

aaa gat ggc gca ttg ggt att ctt gat atc ttt aat tgc gaa tct cgt 624
Lys Asp Gly Ala Leu Gly Ile Leu Asp Ile Phe Asn Cys Glu Ser Arg
195 200 205

gaa ata caa gtt gcc ggg cag aag tac acc ttg agt atc gcc acc gca 672
Glu Ile Gln Val Ala Gly Gln Lys Tyr Thr Leu Ser Ile Ala Thr Ala
210 215 220

aca ttt cac gtt ctc tgg gtg gac gag gcg tgt atg tgg aac ggg gcc 720
Thr Phe His Val Leu Trp Val Asp Glu Ala Cys Met Trp Asn Gly Ala
225 230 235 240

ctg gcc gaa ttt ttt agg gcg ctg cac aat aag ttg ttc ggc gac cgg 768
Leu Ala Glu Phe Phe Arg Ala Leu His Asn Lys Leu Phe Gly Asp Arg
245 250 255

gaa ggc gta gcg cca acg tta acg tac gtg tgt ccg ggg gcc act ccg 816
Glu Gly Val Ala Pro Thr Leu Thr Tyr Val Cys Pro Gly Ala Thr Pro
260 265 270

gag gga acc ccc ttc ccc ccc tac ttt tcc gcg ttt cca cac ctc ccg 864
Glu Gly Thr Pro Phe Pro Pro Tyr Phe Ser Ala Phe Pro His Leu Pro
275 280 285

ctc gtg ttt gga aga ccg cga agg ctc gac gta acc gcg gtc caa gaa 912
Leu Val Phe Gly Arg Pro Arg Arg Leu Asp Val Thr Ala Val Gln Glu
290 295 300

ctc cca aaa gca caa att gcg gta cac tgg ccc ccg ttt aaa gat tca 960
Leu Pro Lys Ala Gln Ile Ala Val His Trp Pro Pro Phe Lys Asp Ser
305 310 315 320

atc tta ggg gat cag ctt ctc ata cct ggc att tca cct aaa aag cca 1008
Ile Leu Gly Asp Gln Leu Leu Ile Pro Gly Ile Ser Pro Lys Lys Pro
325 330 335

ggt acc gta ccc gtt cgt tgg ccg ctt tgg gtg gag gat gtt aac ttg 1056
Gly Thr Val Pro Val Arg Trp Pro Leu Trp Val Glu Asp Val Asn Leu
340 345 350

agt ctc tgc gag acg aca gaa agc gtt gcc cgc ata gtc gac cca cat 1104
Ser Leu Cys Glu Thr Thr Glu Ser Val Ala Arg Ile Val Asp Pro His
355 360 365

tct ata gta atc ata aaa ttt tca tca ctg ttg tgc cag cac cta aaa 1152
Ser Ile Val Ile Ile Lys Phe Ser Ser Leu Leu Cys Gln His Leu Lys
370 375 380

tgc cac cgt gcg ttt gtt aaa aat gag tta gaa tac ata gca acc atc 1200
Cys His Arg Ala Phe Val Lys Asn Glu Leu Glu Tyr Ile Ala Thr Ile
385 390 395 400

tgt tcc agc gac ctt cgc ctc ttc atc caa gag gaa tac aac cgg tta 1248
Cys Ser Ser Asp Leu Arg Leu Phe Ile Gln Glu Glu Tyr Asn Arg Leu
405 410 415

ctt gcc acc att ttt acg tgg gcc gcg gcg agc ggg tat acc tgg gcg 1296
Leu Ala Thr Ile Phe Thr Trp Ala Ala Ala Ser Gly Tyr Thr Trp Ala
420 425 430

gcc att gat aaa aca aca gta ttc atc aag gct ccc cag ctc agc gca 1344
Ala Ile Asp Lys Thr Thr Val Phe Ile Lys Ala Pro Gln Leu Ser Ala
435 440 445

gct gta agt ggt ttc tgc cca tca cta aat agt tgc cgt agg aaa caa 1392
Ala Val Ser Gly Phe Cys Pro Ser Leu Asn Ser Cys Arg Arg Lys Gln
450 455 460

tgt tac gaa ggt taa 1407
Cys Tyr Glu Gly
465

<210> 77
<211> 468
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 77
Met Asn Ala Arg Glu Val Ala Leu Thr Gly His Val Leu His Ile Ser
1 5 10 15
Leu His Ser Thr His Glu Arg Glu Lys Leu Ile Ile Trp Gln Val His
20 25 30
Leu Leu Val Cys Gln Gln Cys Gly Ile Gln Gly Asp Ala Ala Tyr Leu
35 40 45
Phe Val Thr Glu Thr Leu Ser Asn Thr Asp Trp Gly Asn Ile Pro Ala
50 55 60
Ile Asn Arg His Ala Pro Ser Ile Asn Glu His Gly Arg Asn Tyr Met
65 70 75 80
Gln Trp Glu Leu Arg Thr Arg Leu Arg Asn Pro Ile Ile Gln Leu Leu
85 90 95
Ser Arg Gln Pro Gly Ala Val Asn Val Arg Val Ser Glu Pro Asn Met
100 105 110
Val Ile Val Gly Cys Glu Arg Ala Leu Asp His Ser Cys Ser Val Arg
115 120 125
Val Thr Gly Ala Tyr Leu His Cys Asp Thr Thr Met Asp Phe Ser Leu
130 135 140
Asp Ser Val Val Ser Pro Thr Arg Glu Phe Trp Phe Ser Glu Met Phe
145 150 155 160
Ser His Cys Leu Val Ser Asn Ile Glu Val Tyr Leu Lys Thr Thr Gly
165 170 175
Gly Leu Tyr Tyr Arg Ala Ser Ser Ala Thr Gln Cys Arg Lys Arg Ala
180 185 190
Lys Asp Gly Ala Leu Gly Ile Leu Asp Ile Phe Asn Cys Glu Ser Arg

                195   200   205
Glu Ile Gln Val Ala Gly Gln Lys Tyr Thr Leu Ser Ile Ala Thr Ala
        210   215   220
Thr Phe His Val Leu Trp Val Asp Glu Ala Cys Met Trp Asn Gly Ala
225   230   235   240
Leu Ala Glu Phe Phe Arg Ala Leu His Asn Lys Leu Phe Gly Asp Arg
                                        245   250   255
Glu Gly Val Ala Pro Thr Leu Thr Tyr Val Cys Pro Gly Ala Thr Pro
                        260   265   270
Glu Gly Thr Pro Phe Pro Pro Tyr Phe Ser Ala Phe Pro His Leu Pro
                275   280   285
Leu Val Phe Gly Arg Pro Arg Arg Leu Asp Val Thr Ala Val Gln Glu
        290   295   300
Leu Pro Lys Ala Gln Ile Ala Val His Trp Pro Pro Phe Lys Asp Ser
305   310   315   320
Ile Leu Gly Asp Gln Leu Leu Ile Pro Gly Ile Ser Pro Lys Lys Pro
                                        325   330   335
Gly Thr Val Pro Val Arg Trp Pro Leu Trp Val Glu Asp Val Asn Leu
                        340   345   350
Ser Leu Cys Glu Thr Thr Glu Ser Val Ala Arg Ile Val Asp Pro His
                355   360   365
Ser Ile Val Ile Ile Lys Phe Ser Ser Leu Leu Cys Gln His Leu Lys
        370   375   380
Cys His Arg Ala Phe Val Lys Asn Glu Leu Glu Tyr Ile Ala Thr Ile
385   390   395   400
Cys Ser Ser Asp Leu Arg Leu Phe Ile Gln Glu Glu Tyr Asn Arg Leu
                                        405   410   415
Leu Ala Thr Ile Phe Thr Trp Ala Ala Ala Ser Gly Tyr Thr Trp Ala
                        420   425   430
Ala Ile Asp Lys Thr Thr Val Phe Ile Lys Ala Pro Gln Leu Ser Ala
                435   440   445
Ala Val Ser Gly Phe Cys Pro Ser Leu Asn Ser Cys Arg Arg Lys Gln
        450   455   460
Cys Tyr Glu Gly
465

<210> 78
<211> 612
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(612)

<400> 78
atg tta cga agg tta aaa ata aca gtt cat ttc ctt tca cag gaa cag   48
Met Leu Arg Arg Leu Lys Ile Thr Val His Phe Leu Ser Gln Glu Gln
1   5   10   15

caa aag gtc gtg acc cgt ctt gag gcg cat ttg gga ctt ccc gta cag   96
Gln Lys Val Val Thr Arg Leu Glu Ala His Leu Gly Leu Pro Val Gln
            20   25   30

gaa act tcc cac ccg cct gac tgg ctc aag tgt gag gtc tgc tcc gcg   144
Glu Thr Ser His Pro Pro Asp Trp Leu Lys Cys Glu Val Cys Ser Ala
        35   40   45

tcc gtg ttt tta aaa ata cca gcc ggg gtt ttg tat gcc gga ctc gca   192

Ser Val Phe Leu Lys Ile Pro Ala Gly Val Leu Tyr Ala Gly Leu Ala
50 55 60

aga gac ccc acc agg gaa gca aaa cgg gac tcg tgg ctg gac tgt cta 240
Arg Asp Pro Thr Arg Glu Ala Lys Arg Asp Ser Trp Leu Asp Cys Leu
65 70 75 80

gta gaa ggc gcg acg ttg ttg ctt aac aac tca gtg tta ccg att ggg 288
Val Glu Gly Ala Thr Leu Leu Leu Asn Asn Ser Val Leu Pro Ile Gly
85 90 95

gcg ctg gcg ggt atc tta ccc acc ctt ttt gcc aac agg cgg tgt gtt 336
Ala Leu Ala Gly Ile Leu Pro Thr Leu Phe Ala Asn Arg Arg Cys Val
100 105 110

aat ttt tgg ctg ctg cca cgc gcg tgg gta aaa tcg gcg ccc ata tgc 384
Asn Phe Trp Leu Leu Pro Arg Ala Trp Val Lys Ser Ala Pro Ile Cys
115 120 125

cct ccc cta ccg att gac tgt gtt acg cct cca cag ttt gtc gtg aca 432
Pro Pro Leu Pro Ile Asp Cys Val Thr Pro Pro Gln Phe Val Val Thr
130 135 140

aag cgt gga cca atc tgc tgg tac aag gaa tgg ccg tta ccg gtt gac 480
Lys Arg Gly Pro Ile Cys Trp Tyr Lys Glu Trp Pro Leu Pro Val Asp
145 150 155 160

gtt gat ttt atg tac tac cta cag gag gca cta tgt gtt ttt agt gtt 528
Val Asp Phe Met Tyr Tyr Leu Gln Glu Ala Leu Cys Val Phe Ser Val
165 170 175

gtg tcc aac ggg gag ggt acg gag agt cac gcg gac aat ata cga caa 576
Val Ser Asn Gly Glu Gly Thr Glu Ser His Ala Asp Asn Ile Arg Gln
180 185 190

tta gag aag ttt gaa aag gta cta tgt tta ttt taa 612
Leu Glu Lys Phe Glu Lys Val Leu Cys Leu Phe
195 200

<210> 79
<211> 203
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 79
Met Leu Arg Arg Leu Lys Ile Thr Val His Phe Leu Ser Gln Glu Gln
1 5 10 15
Gln Lys Val Val Thr Arg Leu Glu Ala His Leu Gly Leu Pro Val Gln
20 25 30
Glu Thr Ser His Pro Pro Asp Trp Leu Lys Cys Glu Val Cys Ser Ala
35 40 45
Ser Val Phe Leu Lys Ile Pro Ala Gly Val Leu Tyr Ala Gly Leu Ala
50 55 60
Arg Asp Pro Thr Arg Glu Ala Lys Arg Asp Ser Trp Leu Asp Cys Leu
65 70 75 80
Val Glu Gly Ala Thr Leu Leu Leu Asn Asn Ser Val Leu Pro Ile Gly
85 90 95
Ala Leu Ala Gly Ile Leu Pro Thr Leu Phe Ala Asn Arg Arg Cys Val
100 105 110
Asn Phe Trp Leu Leu Pro Arg Ala Trp Val Lys Ser Ala Pro Ile Cys

```
        115                     120                     125
Pro Pro Leu Pro Ile Asp Cys Val Thr Pro Pro Gln Phe Val Val Thr
    130                     135                     140
Lys Arg Gly Pro Ile Cys Trp Tyr Lys Glu Trp Pro Leu Pro Val Asp
145                     150                     155                 160
Val Asp Phe Met Tyr Tyr Leu Gln Glu Ala Leu Cys Val Phe Ser Val
                165                     170                     175
Val Ser Asn Gly Glu Gly Thr Glu Ser His Ala Asp Asn Ile Arg Gln
            180                     185                     190
Leu Glu Lys Phe Glu Lys Val Leu Cys Leu Phe
        195                     200
```

<210> 80
<211> 819
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(819)

<400> 80

```
atg gac caa ata ctg aag cgt cta atg ggg gag cag cac cga tct gag      48
Met Asp Gln Ile Leu Lys Arg Leu Met Gly Glu Gln His Arg Ser Glu
  1               5                  10                  15

gcg gta atg cct gaa aca gaa tgc tcg tcc aga ggg ccc tat aat tat      96
Ala Val Met Pro Glu Thr Glu Cys Ser Ser Arg Gly Pro Tyr Asn Tyr
             20                  25                  30

cca gtt ttt ccg cga ctg atg ttg gaa gta cat aaa aaa aat agc att     144
Pro Val Phe Pro Arg Leu Met Leu Glu Val His Lys Lys Asn Ser Ile
         35                  40                  45

tgc atg gct tcc aat acg cct aag ctg tgt gtt cga gga cga cta aac     192
Cys Met Ala Ser Asn Thr Pro Lys Leu Cys Val Arg Gly Arg Leu Asn
     50                  55                  60

gtt cct gat tta ggc gtg cac gtg cgc acc agg ctt cag tcg gca acg     240
Val Pro Asp Leu Gly Val His Val Arg Thr Arg Leu Gln Ser Ala Thr
 65                  70                  75                  80

ttt act ggg ttt gtc ttt gcg tgc gtg gtc gaa cac gag gac atg atc     288
Phe Thr Gly Phe Val Phe Ala Cys Val Val Glu His Glu Asp Met Ile
                 85                  90                  95

gac gca ctg gac ata tat ccg cac gtt ttt tct gac cgc gtg cag ctg     336
Asp Ala Leu Asp Ile Tyr Pro His Val Phe Ser Asp Arg Val Gln Leu
            100                 105                 110

ttt aag ccg gcg agc gcg agc gta aca gaa ctc tgc tgc att ctt tct     384
Phe Lys Pro Ala Ser Ala Ser Val Thr Glu Leu Cys Cys Ile Leu Ser
        115                 120                 125

atg ctg gaa aac tac gat aag ccc ccc cta tca ttt atc ctg tcc gcg     432
Met Leu Glu Asn Tyr Asp Lys Pro Pro Leu Ser Phe Ile Leu Ser Ala
    130                 135                 140

ctt gac cgg gcc agg tac cta cac gag agg tat acg tgt aat gat tcc     480
```

Leu Asp Arg Ala Arg Tyr Leu His Glu Arg Tyr Thr Cys Asn Asp Ser
145 150 155 160

gcg ttt gta ttg tat ggg atc gaa gtg ata gcg tcg acc ttg gcg gca 528
Ala Phe Val Leu Tyr Gly Ile Glu Val Ile Ala Ser Thr Leu Ala Ala
165 170 175

tac cac gag tta aac cca ccc caa ggc att ttg cgc gtt ccg ccc cta 576
Tyr His Glu Leu Asn Pro Pro Gln Gly Ile Leu Arg Val Pro Pro Leu
180 185 190

gtg cgg ttc aaa ctc cat aag ctg ttg gac gaa aac gca gac gac atg 624
Val Arg Phe Lys Leu His Lys Leu Leu Asp Glu Asn Ala Asp Asp Met
195 200 205

aaa ggc tta ctg aaa ccc att tat ttg gaa tcc ttt cgt cta aca gag 672
Lys Gly Leu Leu Lys Pro Ile Tyr Leu Glu Ser Phe Arg Leu Thr Glu
210 215 220

aac gtt gga gaa gaa gag ggc cac gcg gag act ttt aac ata ttt tat 720
Asn Val Gly Glu Glu Glu Gly His Ala Glu Thr Phe Asn Ile Phe Tyr
225 230 235 240

tgc ggc aca att ttt acc aga cac ctt cat aac gca tcg gta tta aaa 768
Cys Gly Thr Ile Phe Thr Arg His Leu His Asn Ala Ser Val Leu Lys
245 250 255

tat ttt caa ata acc agc cta cac agc atc ccc aga cag act ctg ttt 816
Tyr Phe Gln Ile Thr Ser Leu His Ser Ile Pro Arg Gln Thr Leu Phe
260 265 270

taa 819

<210> 81
<211> 272
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 81
Met Asp Gln Ile Leu Lys Arg Leu Met Gly Glu Gln His Arg Ser Glu
1 5 10 15
Ala Val Met Pro Glu Thr Glu Cys Ser Ser Arg Gly Pro Tyr Asn Tyr
20 25 30
Pro Val Phe Pro Arg Leu Met Leu Glu Val His Lys Lys Asn Ser Ile
35 40 45
Cys Met Ala Ser Asn Thr Pro Lys Leu Cys Val Arg Gly Arg Leu Asn
50 55 60
Val Pro Asp Leu Gly Val His Val Arg Thr Arg Leu Gln Ser Ala Thr
65 70 75 80
Phe Thr Gly Phe Val Phe Ala Cys Val Val Glu His Glu Asp Met Ile
85 90 95
Asp Ala Leu Asp Ile Tyr Pro His Val Phe Ser Asp Arg Val Gln Leu
100 105 110
Phe Lys Pro Ala Ser Ala Ser Val Thr Glu Leu Cys Cys Ile Leu Ser
115 120 125
Met Leu Glu Asn Tyr Asp Lys Pro Pro Leu Ser Phe Ile Leu Ser Ala
130 135 140
Leu Asp Arg Ala Arg Tyr Leu His Glu Arg Tyr Thr Cys Asn Asp Ser
145 150 155 160

Ala Phe Val Leu Tyr Gly Ile Glu Val Ile Ala Ser Thr Leu Ala Ala
165 170 175

Tyr His Glu Leu Asn Pro Pro Gln Gly Ile Leu Arg Val Pro Pro Leu
180 185 190

Val Arg Phe Lys Leu His Lys Leu Leu Asp Glu Asn Ala Asp Asp Met
195 200 205

Lys Gly Leu Leu Lys Pro Ile Tyr Leu Glu Ser Phe Arg Leu Thr Glu
210 215 220

Asn Val Gly Glu Glu Glu Gly His Ala Glu Thr Phe Asn Ile Phe Tyr
225 230 235 240

Cys Gly Thr Ile Phe Thr Arg His Leu His Asn Ala Ser Val Leu Lys
245 250 255

Tyr Phe Gln Ile Thr Ser Leu His Ser Ile Pro Arg Gln Thr Leu Phe
260 265 270

<210> 82
<211> 1731
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1731)

<400> 82
atg ttc aag atg aac ccc gga ttt ggt tcg acg tgt ttg gtc cat ccg 48
Met Phe Lys Met Asn Pro Gly Phe Gly Ser Thr Cys Leu Val His Pro
1 5 10 15

acg gaa ctc tcc atc tcg ctt ttc gag atc cta cag ggc aag tac gcg 96
Thr Glu Leu Ser Ile Ser Leu Phe Glu Ile Leu Gln Gly Lys Tyr Ala
20 25 30

tac gtt cgc ggt caa acg cta cac tcc agt ctt cga aat cca ggg atc 144
Tyr Val Arg Gly Gln Thr Leu His Ser Ser Leu Arg Asn Pro Gly Ile
35 40 45

ttt gga aga cag ctg ttc ata cac ctg tac aag acg gca ttg ggg agc 192
Phe Gly Arg Gln Leu Phe Ile His Leu Tyr Lys Thr Ala Leu Gly Ser
50 55 60

tgc aca tac gac aac gtc tta aag gac tgg acc aac ttc gag acc acc 240
Cys Thr Tyr Asp Asn Val Leu Lys Asp Trp Thr Asn Phe Glu Thr Thr
65 70 75 80

ctg aag acc cgt tgg cgc ggt gtc gag cac ctg acg ccg gaa ttc aaa 288
Leu Lys Thr Arg Trp Arg Gly Val Glu His Leu Thr Pro Glu Phe Lys
85 90 95

cgg tct acg ttt gaa tcg tgg gca cgt acc gtt cgt cta aca gtc gat 336
Arg Ser Thr Phe Glu Ser Trp Ala Arg Thr Val Arg Leu Thr Val Asp
100 105 110

caa ctg tta ctg aac act att aat cag gta cta cac acc agg act gtt 384
Gln Leu Leu Leu Asn Thr Ile Asn Gln Val Leu His Thr Arg Thr Val
115 120 125

tta tcg tac gaa cga tac gtc gac tgg gtg gtc gcc ctt ggg ttg gtg 432
Leu Ser Tyr Glu Arg Tyr Val Asp Trp Val Val Ala Leu Gly Leu Val

130 135 140

cct att gtc agg cgc acc ccc gac ggc gac acg atc gcc agg ata caa 480
Pro Ile Val Arg Arg Thr Pro Asp Gly Asp Thr Ile Ala Arg Ile Gln
145 150 155 160

gca cac tgt caa cag atg agg aaa acg tac gct tcc gga gac gtt acg 528
Ala His Cys Gln Gln Met Arg Lys Thr Tyr Ala Ser Gly Asp Val Thr
165 170 175

ata tcg cgg att gtc gac aaa ctg gca cag gag att acc tcg ata atg 576
Ile Ser Arg Ile Val Asp Lys Leu Ala Gln Glu Ile Thr Ser Ile Met
180 185 190

acc gat gtg acg tct att tat att cca gac tac gcc gag gtg tct gtg 624
Thr Asp Val Thr Ser Ile Tyr Ile Pro Asp Tyr Ala Glu Val Ser Val
195 200 205

gag ttt aac gga gat aag gct gct tat ctc ggc acc tac aga caa aaa 672
Glu Phe Asn Gly Asp Lys Ala Ala Tyr Leu Gly Thr Tyr Arg Gln Lys
210 215 220

gat ata aca gtt gaa gtt gtt tca cgg cct ata att tac aat gga cgc 720
Asp Ile Thr Val Glu Val Val Ser Arg Pro Ile Ile Tyr Asn Gly Arg
225 230 235 240

gtc gcc ttt gat agc cct ctg tac cgg tta ttt acc gcg att atg aca 768
Val Ala Phe Asp Ser Pro Leu Tyr Arg Leu Phe Thr Ala Ile Met Thr
245 250 255

tgt cac agg acc gcg gag cac gca aag tta tgt cag tta cta aat acc 816
Cys His Arg Thr Ala Glu His Ala Lys Leu Cys Gln Leu Leu Asn Thr
260 265 270

gcg cct cta aaa gcg ctg gtc ggt agc acg tgc aac gac atg tat aaa 864
Ala Pro Leu Lys Ala Leu Val Gly Ser Thr Cys Asn Asp Met Tyr Lys
275 280 285

gac ata ttg gct cgc tta gag caa tcg tct caa aaa acc gac ccc aag 912
Asp Ile Leu Ala Arg Leu Glu Gln Ser Ser Gln Lys Thr Asp Pro Lys
290 295 300

agg gag ctc cta aac ttg ctg ata aag ctc gcg gag aac aaa aca gtg 960
Arg Glu Leu Leu Asn Leu Leu Ile Lys Leu Ala Glu Asn Lys Thr Val
305 310 315 320

agc ggc ata aca gac gtt gtt gag gat ttc gtc aca gac gtc tct caa 1008
Ser Gly Ile Thr Asp Val Val Glu Asp Phe Val Thr Asp Val Ser Gln
325 330 335

aac att gtg gac aaa aac aag ctc ttt ggg acc ggt acg gaa agt acg 1056
Asn Ile Val Asp Lys Asn Lys Leu Phe Gly Thr Gly Thr Glu Ser Thr
340 345 350

acc cag ggt ctc cgg aaa cag gtt tca aat acc gtg ttc aaa tgc ctg 1104
Thr Gln Gly Leu Arg Lys Gln Val Ser Asn Thr Val Phe Lys Cys Leu
355 360 365

acg aac cag atc aac gag cag ttc gat aca att tct aac ctt gag aaa 1152
Thr Asn Gln Ile Asn Glu Gln Phe Asp Thr Ile Ser Asn Leu Glu Lys
370 375 380

gaa cga gac gac tac gta aag aag att caa tgc att gaa acg cag ttg 1200
Glu Arg Asp Asp Tyr Val Lys Lys Ile Gln Cys Ile Glu Thr Gln Leu
385 390 395 400

ctt cag agt tta cca gag gga ggc agg ccc aga cac gat att aat ata 1248
Leu Gln Ser Leu Pro Glu Gly Gly Arg Pro Arg His Asp Ile Asn Ile
405 410 415

ctc act cag aat acg ttg cag gct tta tcc ggt ctg cgc gac cca acc 1296
Leu Thr Gln Asn Thr Leu Gln Ala Leu Ser Gly Leu Arg Asp Pro Thr
420 425 430

att aac cta tcc gaa tgt cac atc cct aaa ggc agc tcc gta gta aac 1344
Ile Asn Leu Ser Glu Cys His Ile Pro Lys Gly Ser Ser Val Val Asn
435 440 445

agc ttt ttt tca caa tac gtc cct ccg ttt atg gag atg ctc aaa gag 1392
Ser Phe Phe Ser Gln Tyr Val Pro Pro Phe Met Glu Met Leu Lys Glu
450 455 460

cta act agc ctg tgg gaa ggg gaa atg ttt caa acg tac aac ctc aca 1440
Leu Thr Ser Leu Trp Glu Gly Glu Met Phe Gln Thr Tyr Asn Leu Thr
465 470 475 480

ccc gtg gtc gac aat cag ggg cag cga acg agc atc gcc tac tcg cag 1488
Pro Val Val Asp Asn Gln Gly Gln Arg Thr Ser Ile Ala Tyr Ser Gln
485 490 495

gac acg gtt tcc att ctc tta ggc ccg ttt acg tac ata atc gcc aaa 1536
Asp Thr Val Ser Ile Leu Leu Gly Pro Phe Thr Tyr Ile Ile Ala Lys
500 505 510

ctg act cac atg gac ctc ata aat cac tct ctc att tcc tta agt tta 1584
Leu Thr His Met Asp Leu Ile Asn His Ser Leu Ile Ser Leu Ser Leu
515 520 525

cac gat att gca gac cag ctg tat gtg gac agc agg ctg tct gtg tat 1632
His Asp Ile Ala Asp Gln Leu Tyr Val Asp Ser Arg Leu Ser Val Tyr
530 535 540

ata aat gac ata ggc cac aaa tat tgc gaa caa att agc cag cct gga 1680
Ile Asn Asp Ile Gly His Lys Tyr Cys Glu Gln Ile Ser Gln Pro Gly
545 550 555 560

acc gat gga cca aat act gaa gcg tct aat ggg gga gca gca ccg atc 1728
Thr Asp Gly Pro Asn Thr Glu Ala Ser Asn Gly Gly Ala Ala Pro Ile
565 570 575

tga 1731

<210> 83
<211> 576
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 83
Met Phe Lys Met Asn Pro Gly Phe Gly Ser Thr Cys Leu Val His Pro
1 5 10 15

Thr Glu Leu Ser Ile Ser Leu Phe Glu Ile Leu Gln Gly Lys Tyr Ala
20 25 30
Tyr Val Arg Gly Gln Thr Leu His Ser Ser Leu Arg Asn Pro Gly Ile
35 40 45
Phe Gly Arg Gln Leu Phe Ile His Leu Tyr Lys Thr Ala Leu Gly Ser
50 55 60
Cys Thr Tyr Asp Asn Val Leu Lys Asp Trp Thr Asn Phe Glu Thr Thr
65 70 75 80
Leu Lys Thr Arg Trp Arg Gly Val Glu His Leu Thr Pro Glu Phe Lys
85 90 95
Arg Ser Thr Phe Glu Ser Trp Ala Arg Thr Val Arg Leu Thr Val Asp
100 105 110
Gln Leu Leu Leu Asn Thr Ile Asn Gln Val Leu His Thr Arg Thr Val
115 120 125
Leu Ser Tyr Glu Arg Tyr Val Asp Trp Val Val Ala Leu Gly Leu Val
130 135 140
Pro Ile Val Arg Arg Thr Pro Asp Gly Asp Thr Ile Ala Arg Ile Gln
145 150 155 160
Ala His Cys Gln Gln Met Arg Lys Thr Tyr Ala Ser Gly Asp Val Thr
165 170 175
Ile Ser Arg Ile Val Asp Lys Leu Ala Gln Glu Ile Thr Ser Ile Met
180 185 190
Thr Asp Val Thr Ser Ile Tyr Ile Pro Asp Tyr Ala Glu Val Ser Val
195 200 205
Glu Phe Asn Gly Asp Lys Ala Ala Tyr Leu Gly Thr Tyr Arg Gln Lys
210 215 220
Asp Ile Thr Val Glu Val Val Ser Arg Pro Ile Ile Tyr Asn Gly Arg
225 230 235 240
Val Ala Phe Asp Ser Pro Leu Tyr Arg Leu Phe Thr Ala Ile Met Thr
245 250 255
Cys His Arg Thr Ala Glu His Ala Lys Leu Cys Gln Leu Leu Asn Thr
260 265 270
Ala Pro Leu Lys Ala Leu Val Gly Ser Thr Cys Asn Asp Met Tyr Lys
275 280 285
Asp Ile Leu Ala Arg Leu Glu Gln Ser Ser Gln Lys Thr Asp Pro Lys
290 295 300
Arg Glu Leu Leu Asn Leu Leu Ile Lys Leu Ala Glu Asn Lys Thr Val
305 310 315 320
Ser Gly Ile Thr Asp Val Val Glu Asp Phe Val Thr Asp Val Ser Gln
325 330 335
Asn Ile Val Asp Lys Asn Lys Leu Phe Gly Thr Gly Thr Glu Ser Thr
340 345 350
Thr Gln Gly Leu Arg Lys Gln Val Ser Asn Thr Val Phe Lys Cys Leu
355 360 365
Thr Asn Gln Ile Asn Glu Gln Phe Asp Thr Ile Ser Asn Leu Glu Lys
370 375 380
Glu Arg Asp Asp Tyr Val Lys Lys Ile Gln Cys Ile Glu Thr Gln Leu
385 390 395 400
Leu Gln Ser Leu Pro Glu Gly Gly Arg Pro Arg His Asp Ile Asn Ile
405 410 415
Leu Thr Gln Asn Thr Leu Gln Ala Leu Ser Gly Leu Arg Asp Pro Thr
420 425 430
Ile Asn Leu Ser Glu Cys His Ile Pro Lys Gly Ser Ser Val Val Asn
435 440 445
Ser Phe Phe Ser Gln Tyr Val Pro Pro Phe Met Glu Met Leu Lys Glu
450 455 460
Leu Thr Ser Leu Trp Glu Gly Glu Met Phe Gln Thr Tyr Asn Leu Thr
465 470 475 480
Pro Val Val Asp Asn Gln Gly Gln Arg Thr Ser Ile Ala Tyr Ser Gln
485 490 495
Asp Thr Val Ser Ile Leu Leu Gly Pro Phe Thr Tyr Ile Ile Ala Lys

500 505 510
Leu Thr His Met Asp Leu Ile Asn His Ser Leu Ile Ser Leu Ser Leu
515 520 525
His Asp Ile Ala Asp Gln Leu Tyr Val Asp Ser Arg Leu Ser Val Tyr
530 535 540
Ile Asn Asp Ile Gly His Lys Tyr Cys Glu Gln Ile Ser Gln Pro Gly
545 550 555 560
Thr Asp Gly Pro Asn Thr Glu Ala Ser Asn Gly Gly Ala Ala Pro Ile
565 570 575

<210> 84
<211> 2373
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2373)

<400> 84
atg gag agt tcc gtc gga tgg acc aaa cac gtc gaa cca aat ccg ggg 48
Met Glu Ser Ser Val Gly Trp Thr Lys His Val Glu Pro Asn Pro Gly
1 5 10 15

ttc atc ttg aac atg acg tcc gat gcc aaa gtc agg ggt gtc gtg gat 96
Phe Ile Leu Asn Met Thr Ser Asp Ala Lys Val Arg Gly Val Val Asp
20 25 30

cac gtc agt cgc ctg tca aat ata act acc agc cca ccg gaa atg ggt 144
His Val Ser Arg Leu Ser Asn Ile Thr Thr Ser Pro Pro Glu Met Gly
35 40 45

tgg tac gac ctg gcc ttc gat ccg gct gaa gac tcc ggg ccg ttc ttg 192
Trp Tyr Asp Leu Ala Phe Asp Pro Ala Glu Asp Ser Gly Pro Phe Leu
50 55 60

ccg ttt acc gtt tat cta att acg gga act gct ggt gct ggg aaa agt 240
Pro Phe Thr Val Tyr Leu Ile Thr Gly Thr Ala Gly Ala Gly Lys Ser
65 70 75 80

acc agc ata tcg gcc ctg tac caa aat tta aac tgc ctg atc acg ggc 288
Thr Ser Ile Ser Ala Leu Tyr Gln Asn Leu Asn Cys Leu Ile Thr Gly
85 90 95

gcg acc acc ata gcc gca cag aac cta tcg cgt cgc cta aag acg ttc 336
Ala Thr Thr Ile Ala Ala Gln Asn Leu Ser Arg Arg Leu Lys Thr Phe
100 105 110

tgt ccc acg atc ttc agc gct ttt ggc ttt aag agc cga cac atc aat 384
Cys Pro Thr Ile Phe Ser Ala Phe Gly Phe Lys Ser Arg His Ile Asn
115 120 125

ata gcc gtc aga aaa gct cat cag acc gga gcc gta tcc ata gag caa 432
Ile Ala Val Arg Lys Ala His Gln Thr Gly Ala Val Ser Ile Glu Gln
130 135 140

att cag caa cag gag cta tcg aag tat tgg ccg gtt ata gtg gac att 480
Ile Gln Gln Gln Glu Leu Ser Lys Tyr Trp Pro Val Ile Val Asp Ile
145 150 155 160

```
atg aaa gag gtt atg gcg aaa aaa ccc aat ggc atg tac ggg act ata     528
Met Lys Glu Val Met Ala Lys Lys Pro Asn Gly Met Tyr Gly Thr Ile
                165                 170                 175

tcc aac gcg aat ttt gaa acc ctc tcg aga atg acc gga ccg tgt tta     576
Ser Asn Ala Asn Phe Glu Thr Leu Ser Arg Met Thr Gly Pro Cys Leu
            180                 185                 190

tgg act tcc aat att att gta atc gac gag gcc gga acc ctg tcc tct     624
Trp Thr Ser Asn Ile Ile Val Ile Asp Glu Ala Gly Thr Leu Ser Ser
        195                 200                 205

tac ata ctc acc acc gtc gtg ttc ttt tac tgg ttc cta aac agc tgg     672
Tyr Ile Leu Thr Thr Val Val Phe Phe Tyr Trp Phe Leu Asn Ser Trp
    210                 215                 220

cta aat acc cct ctt tac cgc cag ggg gcg gtt ccg tgc ata gta tgc     720
Leu Asn Thr Pro Leu Tyr Arg Gln Gly Ala Val Pro Cys Ile Val Cys
225                 230                 235                 240

gtc ggt tcg cca acg cag aca aac gcg ttc cag tca act tac aac cac     768
Val Gly Ser Pro Thr Gln Thr Asn Ala Phe Gln Ser Thr Tyr Asn His
                245                 250                 255

ggg acg caa aag acg gag ata tcg tcg tgc gag aac atc cta aca ttc     816
Gly Thr Gln Lys Thr Glu Ile Ser Ser Cys Glu Asn Ile Leu Thr Phe
            260                 265                 270

atg ata ggg aag aag gtc gta tct gag tac gta cac ttg gag aga aac     864
Met Ile Gly Lys Lys Val Val Ser Glu Tyr Val His Leu Glu Arg Asn
        275                 280                 285

tgg gcg ctg ttt ata aac aac aag cgc tgc acc gat ctg cag ttc gga     912
Trp Ala Leu Phe Ile Asn Asn Lys Arg Cys Thr Asp Leu Gln Phe Gly
    290                 295                 300

cac ctt cta aaa att tta gag tat aat ctt ccc att cct gac gaa gtc     960
His Leu Leu Lys Ile Leu Glu Tyr Asn Leu Pro Ile Pro Asp Glu Val
305                 310                 315                 320

atg agt tac gta gac aga ttt gtc gtt cct aaa agt aag att atg gat    1008
Met Ser Tyr Val Asp Arg Phe Val Val Pro Lys Ser Lys Ile Met Asp
                325                 330                 335

cct tta gaa tac att ggc tgg acc cga ctc ttt ttg tca cat agc gag    1056
Pro Leu Glu Tyr Ile Gly Trp Thr Arg Leu Phe Leu Ser His Ser Glu
            340                 345                 350

gta aag gcg tat cta aca aac cta cac aca tgt cta acg cta ggg ggc    1104
Val Lys Ala Tyr Leu Thr Asn Leu His Thr Cys Leu Thr Leu Gly Gly
        355                 360                 365

gat acc agg gac acg aag ctc ttt acc tgt ccc gtg gtg tgc gag gtg    1152
Asp Thr Arg Asp Thr Lys Leu Phe Thr Cys Pro Val Val Cys Glu Val
    370                 375                 380

ttt gtg aag ccg ttt gag gaa tac aag cgg gcc gtc aac ctc acc aac    1200
Phe Val Lys Pro Phe Glu Glu Tyr Lys Arg Ala Val Asn Leu Thr Asn
385                 390                 395                 400
```

ctc acc gtg acc gag tgg gtg aca aaa aac ctc ttt aag tta agt aat 1248
Leu Thr Val Thr Glu Trp Val Thr Lys Asn Leu Phe Lys Leu Ser Asn
405 410 415

tat tcg cag ttt gtg gac cag gac atg tcc ata gtc gcc acg gaa tcc 1296
Tyr Ser Gln Phe Val Asp Gln Asp Met Ser Ile Val Ala Thr Glu Ser
420 425 430

act gaa cgt tcg acc cag gtt acc ttt atc acc aag ttt gtt aaa aac 1344
Thr Glu Arg Ser Thr Gln Val Thr Phe Ile Thr Lys Phe Val Lys Asn
435 440 445

agc cac gtg tct tta aat gga aaa aca aaa aaa tgc ata tgt ggg ttt 1392
Ser His Val Ser Leu Asn Gly Lys Thr Lys Lys Cys Ile Cys Gly Phe
450 455 460

cag ggt acg tac ttt gag ttc aaa aga atc cta gac agc gaa ctc ttc 1440
Gln Gly Thr Tyr Phe Glu Phe Lys Arg Ile Leu Asp Ser Glu Leu Phe
465 470 475 480

gtg gag acg cat tcg caa gac cgt ccg gaa tac gta tac ggt ttc tta 1488
Val Glu Thr His Ser Gln Asp Arg Pro Glu Tyr Val Tyr Gly Phe Leu
485 490 495

aac acg cta ctg tac aac gcc atg tac tcg ttt cac gcg tac ggc gtg 1536
Asn Thr Leu Leu Tyr Asn Ala Met Tyr Ser Phe His Ala Tyr Gly Val
500 505 510

act agg tca cac gag aaa tat ctg caa gac cta aaa ttt gca ccc ctc 1584
Thr Arg Ser His Glu Lys Tyr Leu Gln Asp Leu Lys Phe Ala Pro Leu
515 520 525

ccg gcc gct ctg gca acc ggg cgt gta gac ctt caa acg gtt cgt gaa 1632
Pro Ala Ala Leu Ala Thr Gly Arg Val Asp Leu Gln Thr Val Arg Glu
530 535 540

gag tta aac ctg gaa gac gac atc ttc tac cac gtg tgc agt ccc ccg 1680
Glu Leu Asn Leu Glu Asp Asp Ile Phe Tyr His Val Cys Ser Pro Pro
545 550 555 560

cca ccc gcg ggt atc acc tcc ctc cag gtt ttg gtc gac acg tac tgc 1728
Pro Pro Ala Gly Ile Thr Ser Leu Gln Val Leu Val Asp Thr Tyr Cys
565 570 575

gcc cta aag gac gtg ttc gcc tcc aga ata aag gtg gcg tgt cgc tgg 1776
Ala Leu Lys Asp Val Phe Ala Ser Arg Ile Lys Val Ala Cys Arg Trp
580 585 590

ttt ggc ggg gag ttt gag aag gaa acg ttt tcc gcg ttt acg gtt aac 1824
Phe Gly Gly Glu Phe Glu Lys Glu Thr Phe Ser Ala Phe Thr Val Asn
595 600 605

atg gtc gta agg gac gga gtt gac ttt gtc tcc cct tca gaa cgt ctc 1872
Met Val Val Arg Asp Gly Val Asp Phe Val Ser Pro Ser Glu Arg Leu
610 615 620

aac ggg ctg ttg gcg ttt gca tcg acc gtt gaa tcg tat aaa att aag 1920
Asn Gly Leu Leu Ala Phe Ala Ser Thr Val Glu Ser Tyr Lys Ile Lys
625 630 635 640

ggg tac acg ttt tta ccg gta gcg ttc ggt cgc tgt cag ggt ttg cca 1968

Gly Tyr Thr Phe Leu Pro Val Ala Phe Gly Arg Cys Gln Gly Leu Pro
645 650 655

ctc agt gat gac ctc agg aag aag atg ccc tcc ctg gtc gtg cag gac 2016
Leu Ser Asp Asp Leu Arg Lys Lys Met Pro Ser Leu Val Val Gln Asp
660 665 670

tct agc ggt ttt atc gcg tgc cta gag aat aac ata acc aaa ttg acc 2064
Ser Ser Gly Phe Ile Ala Cys Leu Glu Asn Asn Ile Thr Lys Leu Thr
675 680 685

gaa acc atg gag gac ggg agc gtt ttc caa gtg tgc tgt gcg ggg gac 2112
Glu Thr Met Glu Asp Gly Ser Val Phe Gln Val Cys Cys Ala Gly Asp
690 695 700

tat ggg gtc agc tca aat tta gcc atg acc atc gta aag gca cag gga 2160
Tyr Gly Val Ser Ser Asn Leu Ala Met Thr Ile Val Lys Ala Gln Gly
705 710 715 720

atg tcg ttg gag cga gta gcc gta gta ttt ggg tcc cac aag aac gtc 2208
Met Ser Leu Glu Arg Val Ala Val Val Phe Gly Ser His Lys Asn Val
725 730 735

cag aca agc cac gtg tat gta gca ata tca agg gct gtt aac tca aat 2256
Gln Thr Ser His Val Tyr Val Ala Ile Ser Arg Ala Val Asn Ser Asn
740 745 750

tat ttg gtc atg gac agc aac ccc ctt aaa acc ctc ctc aga gaa cca 2304
Tyr Leu Val Met Asp Ser Asn Pro Leu Lys Thr Leu Leu Arg Glu Pro
755 760 765

gtc gat aac acc tcc gcc aag cat ata gtc cgc gcc ctc cac aac cca 2352
Val Asp Asn Thr Ser Ala Lys His Ile Val Arg Ala Leu His Asn Pro
770 775 780

aac aca acc ctc atc tac taa 2373
Asn Thr Thr Leu Ile Tyr
785 790

<210> 85
<211> 790
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 85
Met Glu Ser Ser Val Gly Trp Thr Lys His Val Glu Pro Asn Pro Gly
1 5 10 15
Phe Ile Leu Asn Met Thr Ser Asp Ala Lys Val Arg Gly Val Val Asp
20 25 30
His Val Ser Arg Leu Ser Asn Ile Thr Thr Ser Pro Pro Glu Met Gly
35 40 45
Trp Tyr Asp Leu Ala Phe Asp Pro Ala Glu Asp Ser Gly Pro Phe Leu
50 55 60
Pro Phe Thr Val Tyr Leu Ile Thr Gly Thr Ala Gly Ala Gly Lys Ser
65 70 75 80
Thr Ser Ile Ser Ala Leu Tyr Gln Asn Leu Asn Cys Leu Ile Thr Gly
85 90 95
Ala Thr Thr Ile Ala Ala Gln Asn Leu Ser Arg Arg Leu Lys Thr Phe
100 105 110
Cys Pro Thr Ile Phe Ser Ala Phe Gly Phe Lys Ser Arg His Ile Asn

115 120 125
Ile Ala Val Arg Lys Ala His Gln Thr Gly Ala Val Ser Ile Glu Gln
130 135 140
Ile Gln Gln Gln Glu Leu Ser Lys Tyr Trp Pro Val Ile Val Asp Ile
145 150 155 160
Met Lys Glu Val Met Ala Lys Lys Pro Asn Gly Met Tyr Gly Thr Ile
165 170 175
Ser Asn Ala Asn Phe Glu Thr Leu Ser Arg Met Thr Gly Pro Cys Leu
180 185 190
Trp Thr Ser Asn Ile Ile Val Ile Asp Glu Ala Gly Thr Leu Ser Ser
195 200 205
Tyr Ile Leu Thr Thr Val Val Phe Phe Tyr Trp Phe Leu Asn Ser Trp
210 215 220
Leu Asn Thr Pro Leu Tyr Arg Gln Gly Ala Val Pro Cys Ile Val Cys
225 230 235 240
Val Gly Ser Pro Thr Gln Thr Asn Ala Phe Gln Ser Thr Tyr Asn His
245 250 255
Gly Thr Gln Lys Thr Glu Ile Ser Ser Cys Glu Asn Ile Leu Thr Phe
260 265 270
Met Ile Gly Lys Lys Val Val Ser Glu Tyr Val His Leu Glu Arg Asn
275 280 285
Trp Ala Leu Phe Ile Asn Asn Lys Arg Cys Thr Asp Leu Gln Phe Gly
290 295 300
His Leu Leu Lys Ile Leu Glu Tyr Asn Leu Pro Ile Pro Asp Glu Val
305 310 315 320
Met Ser Tyr Val Asp Arg Phe Val Val Pro Lys Ser Lys Ile Met Asp
325 330 335
Pro Leu Glu Tyr Ile Gly Trp Thr Arg Leu Phe Leu Ser His Ser Glu
340 345 350
Val Lys Ala Tyr Leu Thr Asn Leu His Thr Cys Leu Thr Leu Gly Gly
355 360 365
Asp Thr Arg Asp Thr Lys Leu Phe Thr Cys Pro Val Val Cys Glu Val
370 375 380
Phe Val Lys Pro Phe Glu Glu Tyr Lys Arg Ala Val Asn Leu Thr Asn
385 390 395 400
Leu Thr Val Thr Glu Trp Val Thr Lys Asn Leu Phe Lys Leu Ser Asn
405 410 415
Tyr Ser Gln Phe Val Asp Gln Asp Met Ser Ile Val Ala Thr Glu Ser
420 425 430
Thr Glu Arg Ser Thr Gln Val Thr Phe Ile Thr Lys Phe Val Lys Asn
435 440 445
Ser His Val Ser Leu Asn Gly Lys Thr Lys Lys Cys Ile Cys Gly Phe
450 455 460
Gln Gly Thr Tyr Phe Glu Phe Lys Arg Ile Leu Asp Ser Glu Leu Phe
465 470 475 480
Val Glu Thr His Ser Gln Asp Arg Pro Glu Tyr Val Tyr Gly Phe Leu
485 490 495
Asn Thr Leu Leu Tyr Asn Ala Met Tyr Ser Phe His Ala Tyr Gly Val
500 505 510
Thr Arg Ser His Glu Lys Tyr Leu Gln Asp Leu Lys Phe Ala Pro Leu
515 520 525
Pro Ala Ala Leu Ala Thr Gly Arg Val Asp Leu Gln Thr Val Arg Glu
530 535 540
Glu Leu Asn Leu Glu Asp Asp Ile Phe Tyr His Val Cys Ser Pro Pro
545 550 555 560
Pro Pro Ala Gly Ile Thr Ser Leu Gln Val Leu Val Asp Thr Tyr Cys
565 570 575
Ala Leu Lys Asp Val Phe Ala Ser Arg Ile Lys Val Ala Cys Arg Trp
580 585 590
Phe Gly Gly Glu Phe Glu Lys Glu Thr Phe Ser Ala Phe Thr Val Asn
595 600 605

Met Val Val Arg Asp Gly Val Asp Phe Val Ser Pro Ser Glu Arg Leu
610 615 620

Asn Gly Leu Leu Ala Phe Ala Ser Thr Val Glu Ser Tyr Lys Ile Lys
625 630 635 640

Gly Tyr Thr Phe Leu Pro Val Ala Phe Gly Arg Cys Gln Gly Leu Pro
645 650 655

Leu Ser Asp Asp Leu Arg Lys Lys Met Pro Ser Leu Val Val Gln Asp
660 665 670

Ser Ser Gly Phe Ile Ala Cys Leu Glu Asn Asn Ile Thr Lys Leu Thr
675 680 685

Glu Thr Met Glu Asp Gly Ser Val Phe Gln Val Cys Cys Ala Gly Asp
690 695 700

Tyr Gly Val Ser Ser Asn Leu Ala Met Thr Ile Val Lys Ala Gln Gly
705 710 715 720

Met Ser Leu Glu Arg Val Ala Val Val Phe Gly Ser His Lys Asn Val
725 730 735

Gln Thr Ser His Val Tyr Val Ala Ile Ser Arg Ala Val Asn Ser Asn
740 745 750

Tyr Leu Val Met Asp Ser Asn Pro Leu Lys Thr Leu Leu Arg Glu Pro
755 760 765

Val Asp Asn Thr Ser Ala Lys His Ile Val Arg Ala Leu His Asn Pro
770 775 780

Asn Thr Thr Leu Ile Tyr
785 790

<210> 86
<211> 1059
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1059)

<400> 86
atg gct atg ttc tta tcc gac ccg tca agg acc ccg ccg gcc act ccc 48
Met Ala Met Phe Leu Ser Asp Pro Ser Arg Thr Pro Pro Ala Thr Pro
1 5 10 15

agg atg ctg ccg att ccc ggg gca cct agg aaa aag aga aca aga cgt 96
Arg Met Leu Pro Ile Pro Gly Ala Pro Arg Lys Lys Arg Thr Arg Arg
20 25 30

ttt ttg ttc gcc ggg tcc cgt acg ggg ttg ccg gta ccg ccc ggg tac 144
Phe Leu Phe Ala Gly Ser Arg Thr Gly Leu Pro Val Pro Pro Gly Tyr
35 40 45

gga ggc ccc cca gtg att gac atg act gca cct gag gat gtt ttt gac 192
Gly Gly Pro Pro Val Ile Asp Met Thr Ala Pro Glu Asp Val Phe Asp
50 55 60

caa gat tcc cct ccc act act cca aaa acg cca gac gaa acc gac tct 240
Gln Asp Ser Pro Pro Thr Thr Pro Lys Thr Pro Asp Glu Thr Asp Ser
65 70 75 80

cat tcg gag aac agc gac tac agt gat gtg gat gaa gaa gac gaa ccg 288
His Ser Glu Asn Ser Asp Tyr Ser Asp Val Asp Glu Glu Asp Glu Pro
85 90 95

```
ccg gtc tcc tcc cca cct aga ata gat ccg cac gca cgg gac ggt gag     336
Pro Val Ser Ser Pro Pro Arg Ile Asp Pro His Ala Arg Asp Gly Glu
            100                 105                 110

tcc ttt aac cag agc ggc cgg cta cct acc gtc att act tcg aca ggt     384
Ser Phe Asn Gln Ser Gly Arg Leu Pro Thr Val Ile Thr Ser Thr Gly
        115                 120                 125

gcc acc aca cca ccg agc gcg ccc gcg cca ctg acc gcg ttt ggt ggt     432
Ala Thr Thr Pro Pro Ser Ala Pro Ala Pro Leu Thr Ala Phe Gly Gly
    130                 135                 140

ccg cgg ccc gtt gcg gtg gtt act ggc caa cac cgc gca ccc cag tca     480
Pro Arg Pro Val Ala Val Val Thr Gly Gln His Arg Ala Pro Gln Ser
145                 150                 155                 160

tct gaa tca gac agc gaa gac gat ttt ttt ata gat gac tac gag gac     528
Ser Glu Ser Asp Ser Glu Asp Asp Phe Phe Ile Asp Asp Tyr Glu Asp
                165                 170                 175

act gat gaa agt ggc gga gaa gct gac ggc ttt tcg cca cgg gcc tcg     576
Thr Asp Glu Ser Gly Gly Glu Ala Asp Gly Phe Ser Pro Arg Ala Ser
            180                 185                 190

ccc gcc tgg tca ggt gat ata tct aga tca ccg gca gag ggc gga tgg     624
Pro Ala Trp Ser Gly Asp Ile Ser Arg Ser Pro Ala Glu Gly Gly Trp
        195                 200                 205

tct tcc gat gaa gaa gaa cct gtt gtt gct gga tca act gca gga caa     672
Ser Ser Asp Glu Glu Glu Pro Val Val Ala Gly Ser Thr Ala Gly Gln
    210                 215                 220

gag acc att att att tca gat gac gac gaa gtg gac gac cga ggc tct     720
Glu Thr Ile Ile Ile Ser Asp Asp Asp Glu Val Asp Asp Arg Gly Ser
225                 230                 235                 240

gtt gag aca tgg gac gaa tca gac gcc gat gag ggc acg ggc gcc act     768
Val Glu Thr Trp Asp Glu Ser Asp Ala Asp Glu Gly Thr Gly Ala Thr
                245                 250                 255

gat gtg att gac ctg tgt tca tct tca gac agc gac gac gat gca gat     816
Asp Val Ile Asp Leu Cys Ser Ser Ser Asp Ser Asp Asp Asp Ala Asp
            260                 265                 270

cac gtt aca agc ggc ggt gtt cgt gcc gcc tgc aag agg cgc gcg tca     864
His Val Thr Ser Gly Gly Val Arg Ala Ala Cys Lys Arg Arg Ala Ser
        275                 280                 285

cgc agg gac tgc aac ggt gat gac gac gtg att tat gtt ggt aca acc     912
Arg Arg Asp Cys Asn Gly Asp Asp Asp Val Ile Tyr Val Gly Thr Thr
    290                 295                 300

cag ggc cct aag aga cgt atg acc tcg acc acg ggg ggc ggg gct act     960
Gln Gly Pro Lys Arg Arg Met Thr Ser Thr Thr Gly Gly Gly Ala Thr
305                 310                 315                 320

tct aac cca gag ggc ccg ggc gtt tct ggt cgt caa act atg gcg gcc     1008
Ser Asn Pro Glu Gly Pro Gly Val Ser Gly Arg Gln Thr Met Ala Ala
                325                 330                 335

acc ccc ccg gtg tgt gga aat gac aac tat ccg tgg ccg tgg ttg gac     1056
```

Thr Pro Pro Val Cys Gly Asn Asp Asn Tyr Pro Trp Pro Trp Leu Asp
340 345 350

tga 1059

<210> 87
<211> 352
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 87
Met Ala Met Phe Leu Ser Asp Pro Ser Arg Thr Pro Pro Ala Thr Pro
1 5 10 15
Arg Met Leu Pro Ile Pro Gly Ala Pro Arg Lys Lys Arg Thr Arg Arg
20 25 30
Phe Leu Phe Ala Gly Ser Arg Thr Gly Leu Pro Val Pro Pro Gly Tyr
35 40 45
Gly Gly Pro Pro Val Ile Asp Met Thr Ala Pro Glu Asp Val Phe Asp
50 55 60
Gln Asp Ser Pro Pro Thr Thr Pro Lys Thr Pro Asp Glu Thr Asp Ser
65 70 75 80
His Ser Glu Asn Ser Asp Tyr Ser Asp Val Asp Glu Glu Asp Glu Pro
85 90 95
Pro Val Ser Ser Pro Pro Arg Ile Asp Pro His Ala Arg Asp Gly Glu
100 105 110
Ser Phe Asn Gln Ser Gly Arg Leu Pro Thr Val Ile Thr Ser Thr Gly
115 120 125
Ala Thr Thr Pro Pro Ser Ala Pro Ala Pro Leu Thr Ala Phe Gly Gly
130 135 140
Pro Arg Pro Val Ala Val Val Thr Gly Gln His Arg Ala Pro Gln Ser
145 150 155 160
Ser Glu Ser Asp Ser Glu Asp Asp Phe Phe Ile Asp Asp Tyr Glu Asp
165 170 175
Thr Asp Glu Ser Gly Gly Glu Ala Asp Gly Phe Ser Pro Arg Ala Ser
180 185 190
Pro Ala Trp Ser Gly Asp Ile Ser Arg Ser Pro Ala Glu Gly Gly Trp
195 200 205
Ser Ser Asp Glu Glu Glu Pro Val Val Ala Gly Ser Thr Ala Gly Gln
210 215 220
Glu Thr Ile Ile Ile Ser Asp Asp Asp Glu Val Asp Asp Arg Gly Ser
225 230 235 240
Val Glu Thr Trp Asp Glu Ser Asp Ala Asp Glu Gly Thr Gly Ala Thr
245 250 255
Asp Val Ile Asp Leu Cys Ser Ser Ser Asp Ser Asp Asp Asp Ala Asp
260 265 270
His Val Thr Ser Gly Gly Val Arg Ala Ala Cys Lys Arg Arg Ala Ser
275 280 285
Arg Arg Asp Cys Asn Gly Asp Asp Asp Val Ile Tyr Val Gly Thr Thr
290 295 300
Gln Gly Pro Lys Arg Arg Met Thr Ser Thr Thr Gly Gly Gly Ala Thr
305 310 315 320
Ser Asn Pro Glu Gly Pro Gly Val Ser Gly Arg Gln Thr Met Ala Ala
325 330 335
Thr Pro Pro Val Cys Gly Asn Asp Asn Tyr Pro Trp Pro Trp Leu Asp
340 345 350

<210> 88

<211> 768
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(768)

<400> 88
atg gag ggt tgg ctt aaa aca ata gtt tgg agt aaa atg tca cca gaa 48
Met Glu Gly Trp Leu Lys Thr Ile Val Trp Ser Lys Met Ser Pro Glu
1 5 10 15

gtg tta gaa gaa ccc tcc aca caa acg ttg cta ctt agc gac tca tgg 96
Val Leu Glu Glu Pro Ser Thr Gln Thr Leu Leu Leu Ser Asp Ser Trp
20 25 30

ctt gaa ttt tta aat ctt tct cca ttt tta aaa cag aag ctc gcg gct 144
Leu Glu Phe Leu Asn Leu Ser Pro Phe Leu Lys Gln Lys Leu Ala Ala
35 40 45

ctg ctg aag cgc gtt atg gat atg agc aac gta acc gtg att tat cca 192
Leu Leu Lys Arg Val Met Asp Met Ser Asn Val Thr Val Ile Tyr Pro
50 55 60

ccg ata gat aga att atg tgg tgg tcg tat tgt tgc gaa ccg gag gat 240
Pro Ile Asp Arg Ile Met Trp Trp Ser Tyr Cys Cys Glu Pro Glu Asp
65 70 75 80

att aaa gtc gtg atc ctt ggc cag gat cct tac cat cgc ggt caa gcc 288
Ile Lys Val Val Ile Leu Gly Gln Asp Pro Tyr His Arg Gly Gln Ala
85 90 95

acc gga cta gcc ttt agc gtt gct ccg gac tac agt ata cct cca agc 336
Thr Gly Leu Ala Phe Ser Val Ala Pro Asp Tyr Ser Ile Pro Pro Ser
100 105 110

ctc aaa aat att ttt aaa gag ata gcc aat act gta cct ggg ttc acc 384
Leu Lys Asn Ile Phe Lys Glu Ile Ala Asn Thr Val Pro Gly Phe Thr
115 120 125

gct cct tct cac ggg tgc tta gac tgt tgg gca aaa cgg gga gtt ctg 432
Ala Pro Ser His Gly Cys Leu Asp Cys Trp Ala Lys Arg Gly Val Leu
130 135 140

ctt tta aac acc att ctg acg gtg gaa aga ggg aag gcg ggg tca cac 480
Leu Leu Asn Thr Ile Leu Thr Val Glu Arg Gly Lys Ala Gly Ser His
145 150 155 160

gcc aac ctt ggc tgg gat tgg ttc acc agc tac ata ata agc tgc ctt 528
Ala Asn Leu Gly Trp Asp Trp Phe Thr Ser Tyr Ile Ile Ser Cys Leu
165 170 175

tct gcc aag ctt cag agg tgc gtt ttt atg ctg tgg gga aga aag gct 576
Ser Ala Lys Leu Gln Arg Cys Val Phe Met Leu Trp Gly Arg Lys Ala
180 185 190

ata gac aag gcg acg ctg ata aac gga cag aga cat ctc gtc ctc aag 624
Ile Asp Lys Ala Thr Leu Ile Asn Gly Gln Arg His Leu Val Leu Lys
195 200 205

gcg cgc cat ccg tct ccc ctg gct acc gcc cac gcg gca acc ggt tct 672
Ala Arg His Pro Ser Pro Leu Ala Thr Ala His Ala Ala Thr Gly Ser
210 215 220

ccc tgg ccg cag ttt cta ggc tgc aac cat ttt aaa ctt gcg aat gac 720
Pro Trp Pro Gln Phe Leu Gly Cys Asn His Phe Lys Leu Ala Asn Asp
225 230 235 240

tat cta gtt cag aat cgg agg ggt gcc gtt gac tgg aat att aac tga 768
Tyr Leu Val Gln Asn Arg Arg Gly Ala Val Asp Trp Asn Ile Asn
245 250 255

<210> 89
<211> 255
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 89
Met Glu Gly Trp Leu Lys Thr Ile Val Trp Ser Lys Met Ser Pro Glu
1 5 10 15
Val Leu Glu Glu Pro Ser Thr Gln Thr Leu Leu Leu Ser Asp Ser Trp
20 25 30
Leu Glu Phe Leu Asn Leu Ser Pro Phe Leu Lys Gln Lys Leu Ala Ala
35 40 45
Leu Leu Lys Arg Val Met Asp Met Ser Asn Val Thr Val Ile Tyr Pro
50 55 60
Pro Ile Asp Arg Ile Met Trp Trp Ser Tyr Cys Cys Glu Pro Glu Asp
65 70 75 80
Ile Lys Val Val Ile Leu Gly Gln Asp Pro Tyr His Arg Gly Gln Ala
85 90 95
Thr Gly Leu Ala Phe Ser Val Ala Pro Asp Tyr Ser Ile Pro Pro Ser
100 105 110
Leu Lys Asn Ile Phe Lys Glu Ile Ala Asn Thr Val Pro Gly Phe Thr
115 120 125
Ala Pro Ser His Gly Cys Leu Asp Cys Trp Ala Lys Arg Gly Val Leu
130 135 140
Leu Leu Asn Thr Ile Leu Thr Val Glu Arg Gly Lys Ala Gly Ser His
145 150 155 160
Ala Asn Leu Gly Trp Asp Trp Phe Thr Ser Tyr Ile Ile Ser Cys Leu
165 170 175
Ser Ala Lys Leu Gln Arg Cys Val Phe Met Leu Trp Gly Arg Lys Ala
180 185 190
Ile Asp Lys Ala Thr Leu Ile Asn Gly Gln Arg His Leu Val Leu Lys
195 200 205
Ala Arg His Pro Ser Pro Leu Ala Thr Ala His Ala Ala Thr Gly Ser
210 215 220
Pro Trp Pro Gln Phe Leu Gly Cys Asn His Phe Lys Leu Ala Asn Asp
225 230 235 240
Tyr Leu Val Gln Asn Arg Arg Gly Ala Val Asp Trp Asn Ile Asn
245 250 255

<210> 90
<211> 510
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS

<222> (1)..(510)

<400> 90
atg aga agt atg tat acc tta agt tta ttt att aca tgt gga ttt ttt 48
Met Arg Ser Met Tyr Thr Leu Ser Leu Phe Ile Thr Cys Gly Phe Phe
1 5 10 15

tta att aca tgt tgt acc ggg ctt gta gta aat cca tgt tgt aaa ata 96
Leu Ile Thr Cys Cys Thr Gly Leu Val Val Asn Pro Cys Cys Lys Ile
20 25 30

ata cca ctt agt gac ttt ata ttt ccg gag ccg ttt gag att gct tct 144
Ile Pro Leu Ser Asp Phe Ile Phe Pro Glu Pro Phe Glu Ile Ala Ser
35 40 45

ttt cat tta aca aat tta gct tta tgt cca ggg ctg tgt aca gca aca 192
Phe His Leu Thr Asn Leu Ala Leu Cys Pro Gly Leu Cys Thr Ala Thr
50 55 60

ttg cga tat aaa gct gat cgc tct acc aca gaa ata tgt gta aat gga 240
Leu Arg Tyr Lys Ala Asp Arg Ser Thr Thr Glu Ile Cys Val Asn Gly
65 70 75 80

ttt cac ctt agg gcg ttt ttt ata agg att ttg tac aag tta aat tac 288
Phe His Leu Arg Ala Phe Phe Ile Arg Ile Leu Tyr Lys Leu Asn Tyr
85 90 95

agt gtc ccc aga gaa gag ctt caa ttg ttg aat tat atg caa tat agt 336
Ser Val Pro Arg Glu Glu Leu Gln Leu Leu Asn Tyr Met Gln Tyr Ser
100 105 110

ctt gat gag ttt tta gcc gag ttt gaa gac ttt cac ata aac ggt agt 384
Leu Asp Glu Phe Leu Ala Glu Phe Glu Asp Phe His Ile Asn Gly Ser
115 120 125

gaa agc gga act gct tat acg cgg cca cct ctg ttg gat ttc tca gac 432
Glu Ser Gly Thr Ala Tyr Thr Arg Pro Pro Leu Leu Asp Phe Ser Asp
130 135 140

aga agt aca aaa gtt tca cgt ata cgt aaa gta att acc aga cgc ggg 480
Arg Ser Thr Lys Val Ser Arg Ile Arg Lys Val Ile Thr Arg Arg Gly
145 150 155 160

gat tta tgg agg gtt ggc tta aaa caa tag 510
Asp Leu Trp Arg Val Gly Leu Lys Gln
165 170

<210> 91
<211> 169
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 91
Met Arg Ser Met Tyr Thr Leu Ser Leu Phe Ile Thr Cys Gly Phe Phe
1 5 10 15
Leu Ile Thr Cys Cys Thr Gly Leu Val Val Asn Pro Cys Cys Lys Ile
20 25 30
Ile Pro Leu Ser Asp Phe Ile Phe Pro Glu Pro Phe Glu Ile Ala Ser
35 40 45
Phe His Leu Thr Asn Leu Ala Leu Cys Pro Gly Leu Cys Thr Ala Thr

```
    50                  55                  60
Leu Arg Tyr Lys Ala Asp Arg Ser Thr Thr Glu Ile Cys Val Asn Gly
65                  70                  75                  80
Phe His Leu Arg Ala Phe Phe Ile Arg Ile Leu Tyr Lys Leu Asn Tyr
                85                  90                  95
Ser Val Pro Arg Glu Glu Leu Gln Leu Leu Asn Tyr Met Gln Tyr Ser
            100                 105                 110
Leu Asp Glu Phe Leu Ala Glu Phe Glu Asp Phe His Ile Asn Gly Ser
        115                 120                 125
Glu Ser Gly Thr Ala Tyr Thr Arg Pro Pro Leu Leu Asp Phe Ser Asp
    130                 135                 140
Arg Ser Thr Lys Val Ser Arg Ile Arg Lys Val Ile Thr Arg Arg Gly
145                 150                 155                 160
Asp Leu Trp Arg Val Gly Leu Lys Gln
                165
```

```
<210> 92
<211> 1170
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1170)

<400> 92
atg gct gtt tcc att cca gtt aag ggt gtt aac agg gag aca gag tct      48
Met Ala Val Ser Ile Pro Val Lys Gly Val Asn Arg Glu Thr Glu Ser
1               5                   10                  15

aac tgg aga tcc atc gtg act acc ttt gaa cga cat gga aac gcg gat      96
Asn Trp Arg Ser Ile Val Thr Thr Phe Glu Arg His Gly Asn Ala Asp
            20                  25                  30

aga gcc atc cgg tcg ctt ttg cga ttt ttt aaa ggt gtc gac cat cct     144
Arg Ala Ile Arg Ser Leu Leu Arg Phe Phe Lys Gly Val Asp His Pro
        35                  40                  45

ggg ttt tta gcg tct ttg gtt ata ctt aaa gat gta act ata gat tca     192
Gly Phe Leu Ala Ser Leu Val Ile Leu Lys Asp Val Thr Ile Asp Ser
    50                  55                  60

gag aaa acg ata gaa aga acg gac ctt att ccc ctt tta cag gga gtg     240
Glu Lys Thr Ile Glu Arg Thr Asp Leu Ile Pro Leu Leu Gln Gly Val
65                  70                  75                  80

agg ttt gtt acc caa cag ata tat atg cac cta aaa gat cac gcg agc     288
Arg Phe Val Thr Gln Gln Ile Tyr Met His Leu Lys Asp His Ala Ser
                85                  90                  95

gaa tct cct atg gcc gag ata tgg cgc gac tgc aaa gag cgc ttc tgc     336
Glu Ser Pro Met Ala Glu Ile Trp Arg Asp Cys Lys Glu Arg Phe Cys
            100                 105                 110

ctc gct ctg gaa ctg gcg tgc ggc tgt caa agg tgc gcg agc gcc gcc     384
Leu Ala Leu Glu Leu Ala Cys Gly Cys Gln Arg Cys Ala Ser Ala Ala
        115                 120                 125

agg cag cta cgg gcc tgt cag caa gcc tgc agg cca cct aag ctg aat     432
```

```
Arg Gln Leu Arg Ala Cys Gln Gln Ala Cys Arg Pro Pro Lys Leu Asn
    130                 135                 140

cct cac aag cag cag tgt gtg gcc gct cgt ctc tta acc gca gtt tat     480
Pro His Lys Gln Gln Cys Val Ala Ala Arg Leu Leu Thr Ala Val Tyr
145                 150                 155                 160

aac cag atg gtg ctc aga act agg gtt tcg gtt tcg gaa ttt tgt ctg     528
Asn Gln Met Val Leu Arg Thr Arg Val Ser Val Ser Glu Phe Cys Leu
                165                 170                 175

aac gct ttg atg tgc gtt cct aga gag ttc ggg ttt gtg agc ggg gat     576
Asn Ala Leu Met Cys Val Pro Arg Glu Phe Gly Phe Val Ser Gly Asp
            180                 185                 190

gtc cgg gtg gaa acc agc cgc gtt gcc tcc tgt ttg aac ttg tcg tgg     624
Val Arg Val Glu Thr Ser Arg Val Ala Ser Cys Leu Asn Leu Ser Trp
        195                 200                 205

tta tac ttg att tta gac tcg tat gtt cga aca gat tta aca aat ctg     672
Leu Tyr Leu Ile Leu Asp Ser Tyr Val Arg Thr Asp Leu Thr Asn Leu
    210                 215                 220

gaa atg gca atg agc cgt gcc tgc cgc att cac ggc ctt agc gcc ggg     720
Glu Met Ala Met Ser Arg Ala Cys Arg Ile His Gly Leu Ser Ala Gly
225                 230                 235                 240

gac ccg ttt tat tcc gcc ctc gtg tgg tta aaa aat agt tac gca tgt     768
Asp Pro Phe Tyr Ser Ala Leu Val Trp Leu Lys Asn Ser Tyr Ala Cys
                245                 250                 255

gac acg aat aca ttt ttt ttc acc gtc aat tca acc agt gtc acg act     816
Asp Thr Asn Thr Phe Phe Phe Thr Val Asn Ser Thr Ser Val Thr Thr
            260                 265                 270

cca atc tta atg gac ata tgt gcg tca ttg acg ggt ccg gtt cct gat     864
Pro Ile Leu Met Asp Ile Cys Ala Ser Leu Thr Gly Pro Val Pro Asp
        275                 280                 285

gtt ata aaa ata aat atg ctg ccc ctc gtg aac gat caa atg cat cct     912
Val Ile Lys Ile Asn Met Leu Pro Leu Val Asn Asp Gln Met His Pro
    290                 295                 300

tcc gtt tgt gtg gag cgt gcc aat ttt acg ggg tct tgt ccc aaa gtg     960
Ser Val Cys Val Glu Arg Ala Asn Phe Thr Gly Ser Cys Pro Lys Val
305                 310                 315                 320

tca ccc act cac cat ttg gat ggg cta aaa tta gaa aca aca tcc ttg     1008
Ser Pro Thr His His Leu Asp Gly Leu Lys Leu Glu Thr Thr Ser Leu
                325                 330                 335

aca tta gca gct gat tca ctg gat gat att tta caa gcc ctg gaa tta     1056
Thr Leu Ala Ala Asp Ser Leu Asp Asp Ile Leu Gln Ala Leu Glu Leu
            340                 345                 350

att tgt ggc gat gac gag ggt atc ctg gac tca tat att tcc gac att     1104
Ile Cys Gly Asp Asp Glu Gly Ile Leu Asp Ser Tyr Ile Ser Asp Ile
        355                 360                 365

aac acg gaa acc gag gtt gat gag tca tca att gaa gag gaa att gtc     1152
Asn Thr Glu Thr Glu Val Asp Glu Ser Ser Ile Glu Glu Glu Ile Val
```

    370   375   380

ttt gag gag ctt agc taa   1170
Phe Glu Glu Leu Ser
385   390

<210> 93
<211> 389
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 93
Met Ala Val Ser Ile Pro Val Lys Gly Val Asn Arg Glu Thr Glu Ser
1   5   10   15
Asn Trp Arg Ser Ile Val Thr Thr Phe Glu Arg His Gly Asn Ala Asp
            20   25   30
Arg Ala Ile Arg Ser Leu Leu Arg Phe Phe Lys Gly Val Asp His Pro
        35   40   45
Gly Phe Leu Ala Ser Leu Val Ile Leu Lys Asp Val Thr Ile Asp Ser
    50   55   60
Glu Lys Thr Ile Glu Arg Thr Asp Leu Ile Pro Leu Leu Gln Gly Val
65   70   75   80
Arg Phe Val Thr Gln Gln Ile Tyr Met His Leu Lys Asp His Ala Ser
                85   90   95
Glu Ser Pro Met Ala Glu Ile Trp Arg Asp Cys Lys Glu Arg Phe Cys
            100   105   110
Leu Ala Leu Glu Leu Ala Cys Gly Cys Gln Arg Cys Ala Ser Ala Ala
        115   120   125
Arg Gln Leu Arg Ala Cys Gln Gln Ala Cys Arg Pro Pro Lys Leu Asn
    130   135   140
Pro His Lys Gln Gln Cys Val Ala Ala Arg Leu Leu Thr Ala Val Tyr
145   150   155   160
Asn Gln Met Val Leu Arg Thr Arg Val Ser Val Ser Glu Phe Cys Leu
                165   170   175
Asn Ala Leu Met Cys Val Pro Arg Glu Phe Gly Phe Val Ser Gly Asp
            180   185   190
Val Arg Val Glu Thr Ser Arg Val Ala Ser Cys Leu Asn Leu Ser Trp
        195   200   205
Leu Tyr Leu Ile Leu Asp Ser Tyr Val Arg Thr Asp Leu Thr Asn Leu
    210   215   220
Glu Met Ala Met Ser Arg Ala Cys Arg Ile His Gly Leu Ser Ala Gly
225   230   235   240
Asp Pro Phe Tyr Ser Ala Leu Val Trp Leu Lys Asn Ser Tyr Ala Cys
                245   250   255
Asp Thr Asn Thr Phe Phe Phe Thr Val Asn Ser Thr Ser Val Thr Thr
            260   265   270
Pro Ile Leu Met Asp Ile Cys Ala Ser Leu Thr Gly Pro Val Pro Asp
        275   280   285
Val Ile Lys Ile Asn Met Leu Pro Leu Val Asn Asp Gln Met His Pro
    290   295   300
Ser Val Cys Val Glu Arg Ala Asn Phe Thr Gly Ser Cys Pro Lys Val
305   310   315   320
Ser Pro Thr His His Leu Asp Gly Leu Lys Leu Glu Thr Thr Ser Leu
                325   330   335
Thr Leu Ala Ala Asp Ser Leu Asp Asp Ile Leu Gln Ala Leu Glu Leu
            340   345   350
Ile Cys Gly Asp Asp Glu Gly Ile Leu Asp Ser Tyr Ile Ser Asp Ile
        355   360   365
Asn Thr Glu Thr Glu Val Asp Glu Ser Ser Ile Glu Glu Glu Ile Val
    370   375   380

Phe Glu Glu Leu Ser
385

<210> 94
<211> 906
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(906)

<400> 94
atg agt cgc cat tac ggg aaa gac cat ctt tta aat cat atg tac aag 48
Met Ser Arg His Tyr Gly Lys Asp His Leu Leu Asn His Met Tyr Lys
1 5 10 15

ttc cac tat ccg cca ctc ggc atg ata gtc ggt gaa atg aac aca ctg 96
Phe His Tyr Pro Pro Leu Gly Met Ile Val Gly Glu Met Asn Thr Leu
20 25 30

acc gtc aat gcc cgc aat ccg ctc tac caa gcg gca acc ctt aga gtg 144
Thr Val Asn Ala Arg Asn Pro Leu Tyr Gln Ala Ala Thr Leu Arg Val
35 40 45

gaa cgg gcc cta tac ctg tct aag ata ctc cag gtg ctc atg caa cac 192
Glu Arg Ala Leu Tyr Leu Ser Lys Ile Leu Gln Val Leu Met Gln His
50 55 60

aga caa gga gag cgg ttc atc gta ccg cag tgt cgg agc aac atg gtt 240
Arg Gln Gly Glu Arg Phe Ile Val Pro Gln Cys Arg Ser Asn Met Val
65 70 75 80

tac tgt cta aag gag ctc cac aaa att acg aac gat aga att cgg gga 288
Tyr Cys Leu Lys Glu Leu His Lys Ile Thr Asn Asp Arg Ile Arg Gly
85 90 95

ctt att aac tca gta ctg ccc ctg gtg gat gcg ggc tgc gta gga ttt 336
Leu Ile Asn Ser Val Leu Pro Leu Val Asp Ala Gly Cys Val Gly Phe
100 105 110

gac gaa gaa ctc gta cgc gtc ctg ccg gag atc cta aaa ctg gag tac 384
Asp Glu Glu Leu Val Arg Val Leu Pro Glu Ile Leu Lys Leu Glu Tyr
115 120 125

ccc cac gcc cac gaa ttg tta ccc cct cac gat ccg acc tcc ccc ctg 432
Pro His Ala His Glu Leu Leu Pro Pro His Asp Pro Thr Ser Pro Leu
130 135 140

agc tgg tgc cta tct cac atg gtc ggt gta acc aaa acc ttc aaa ggg 480
Ser Trp Cys Leu Ser His Met Val Gly Val Thr Lys Thr Phe Lys Gly
145 150 155 160

gag gtg aag gag atg ata gac act ttt cac gac ctg tcc gta cca tcg 528
Glu Val Lys Glu Met Ile Asp Thr Phe His Asp Leu Ser Val Pro Ser
165 170 175

ttt cag tac ctg gcg agt ttg gtt aaa aaa ttt ttt ctt gtt gag gag 576
Phe Gln Tyr Leu Ala Ser Leu Val Lys Lys Phe Phe Leu Val Glu Glu

180 185 190

gtt ata tac gag gac tac cag gac acg cag ttt aac gtg ttt tta aat 624
Val Ile Tyr Glu Asp Tyr Gln Asp Thr Gln Phe Asn Val Phe Leu Asn
195 200 205

ctt tgt ttt ttt tgg acc act gtc ata aag atg tac cag agt tgc att 672
Leu Cys Phe Phe Trp Thr Thr Val Ile Lys Met Tyr Gln Ser Cys Ile
210 215 220

ttt aaa gac aag cta ttg gac acg att aaa gct tgc ata gag ctt cta 720
Phe Lys Asp Lys Leu Leu Asp Thr Ile Lys Ala Cys Ile Glu Leu Leu
225 230 235 240

aaa ggc gag gcc agg cag ttt ttt ggt tgg tac gac cta aac acg cca 768
Lys Gly Glu Ala Arg Gln Phe Phe Gly Trp Tyr Asp Leu Asn Thr Pro
245 250 255

aat tta ggt tca tcg gca cta gta aag tac aca gag cac ctg atc cga 816
Asn Leu Gly Ser Ser Ala Leu Val Lys Tyr Thr Glu His Leu Ile Arg
260 265 270

gca cta agt gtg gat tca tca gcc att ccc att ggc gag ata tgc tcc 864
Ala Leu Ser Val Asp Ser Ser Ala Ile Pro Ile Gly Glu Ile Cys Ser
275 280 285

cac cta cac cac tgt aaa cac gcc ctc ctg aat ctt gaa taa 906
His Leu His His Cys Lys His Ala Leu Leu Asn Leu Glu
290 295 300

<210> 95
<211> 301
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 95
Met Ser Arg His Tyr Gly Lys Asp His Leu Leu Asn His Met Tyr Lys
1 5 10 15
Phe His Tyr Pro Pro Leu Gly Met Ile Val Gly Glu Met Asn Thr Leu
20 25 30
Thr Val Asn Ala Arg Asn Pro Leu Tyr Gln Ala Ala Thr Leu Arg Val
35 40 45
Glu Arg Ala Leu Tyr Leu Ser Lys Ile Leu Gln Val Leu Met Gln His
50 55 60
Arg Gln Gly Glu Arg Phe Ile Val Pro Gln Cys Arg Ser Asn Met Val
65 70 75 80
Tyr Cys Leu Lys Glu Leu His Lys Ile Thr Asn Asp Arg Ile Arg Gly
85 90 95
Leu Ile Asn Ser Val Leu Pro Leu Val Asp Ala Gly Cys Val Gly Phe
100 105 110
Asp Glu Glu Leu Val Arg Val Leu Pro Glu Ile Leu Lys Leu Glu Tyr
115 120 125
Pro His Ala His Glu Leu Leu Pro Pro His Asp Pro Thr Ser Pro Leu
130 135 140
Ser Trp Cys Leu Ser His Met Val Gly Val Thr Lys Thr Phe Lys Gly
145 150 155 160
Glu Val Lys Glu Met Ile Asp Thr Phe His Asp Leu Ser Val Pro Ser
165 170 175
Phe Gln Tyr Leu Ala Ser Leu Val Lys Lys Phe Phe Leu Val Glu Glu
180 185 190

Val Ile Tyr Glu Asp Tyr Gln Asp Thr Gln Phe Asn Val Phe Leu Asn
195 200 205

Leu Cys Phe Phe Trp Thr Thr Val Ile Lys Met Tyr Gln Ser Cys Ile
210 215 220

Phe Lys Asp Lys Leu Leu Asp Thr Ile Lys Ala Cys Ile Glu Leu Leu
225 230 235 240

Lys Gly Glu Ala Arg Gln Phe Phe Gly Trp Tyr Asp Leu Asn Thr Pro
245 250 255

Asn Leu Gly Ser Ser Ala Leu Val Lys Tyr Thr Glu His Leu Ile Arg
260 265 270

Ala Leu Ser Val Asp Ser Ser Ala Ile Pro Ile Gly Glu Ile Cys Ser
275 280 285

His Leu His His Cys Lys His Ala Leu Leu Asn Leu Glu
290 295 300

<210> 96
<211> 1545
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1545)

<400> 96
atg gaa tgt gcc tcc ctg ggt ccc atc tcc ggg ctc atc gcg gat ctg 48
Met Glu Cys Ala Ser Leu Gly Pro Ile Ser Gly Leu Ile Ala Asp Leu
1 5 10 15

aat ctc ttt aac ctg ttc tgc ctc tac cgg ggt tcc cgg gtt aaa acc 96
Asn Leu Phe Asn Leu Phe Cys Leu Tyr Arg Gly Ser Arg Val Lys Thr
20 25 30

cgc ggg gcc gcc acc tgt aac gtc ccg tgc gca gaa tgc gcc cag ggc 144
Arg Gly Ala Ala Thr Cys Asn Val Pro Cys Ala Glu Cys Ala Gln Gly
35 40 45

atc gtg aga att ctc acc gag agg gcc ctt tgt tgt acc gaa aag atg 192
Ile Val Arg Ile Leu Thr Glu Arg Ala Leu Cys Cys Thr Glu Lys Met
50 55 60

ttc ata gcg tct gcc tgc agc ggg gtc gtt ata ccg ccc cag ttg gcc 240
Phe Ile Ala Ser Ala Cys Ser Gly Val Val Ile Pro Pro Gln Leu Ala
65 70 75 80

aga gtt ctg cac gac gtc tac gcg gaa atg aag gcc aag tgt ttg ggg 288
Arg Val Leu His Asp Val Tyr Ala Glu Met Lys Ala Lys Cys Leu Gly
85 90 95

gcg tgg cgt cgt ctc ata tgt tgc agg cgg cct atc atg gct att gcc 336
Ala Trp Arg Arg Leu Ile Cys Cys Arg Arg Pro Ile Met Ala Ile Ala
100 105 110

gac tcg gtc ctc gtc act tat aat acc ctg gat gcc gag gga aaa ctg 384
Asp Ser Val Leu Val Thr Tyr Asn Thr Leu Asp Ala Glu Gly Lys Leu
115 120 125

gaa ctt aga ctc aaa gca ctg tgc aaa ctt gtt ttt caa ccc atc ttt 432
Glu Leu Arg Leu Lys Ala Leu Cys Lys Leu Val Phe Gln Pro Ile Phe

```
        130                 135                 140

ctt caa aga atc tta gcc cct atg cag cta ctg gcc aac ggg aag atg    480
Leu Gln Arg Ile Leu Ala Pro Met Gln Leu Leu Ala Asn Gly Lys Met
145                 150                 155                 160

gtt cct gac aat tat ttt acc atc acc ggt acg gcc gag aag agg cgc    528
Val Pro Asp Asn Tyr Phe Thr Ile Thr Gly Thr Ala Glu Lys Arg Arg
                165                 170                 175

cct gtc gtg act gga agt act agc ggg atg acg tgt ccg gga agc agc    576
Pro Val Val Thr Gly Ser Thr Ser Gly Met Thr Cys Pro Gly Ser Ser
            180                 185                 190

ctt gtc ccc gat tcc tta atc ctg cca gta tgc gaa ccg ggg ctt ctc    624
Leu Val Pro Asp Ser Leu Ile Leu Pro Val Cys Glu Pro Gly Leu Leu
        195                 200                 205

ccg gca ccc ctg gtt gac ctc agt aat gtc tta gaa aat cca gaa atc    672
Pro Ala Pro Leu Val Asp Leu Ser Asn Val Leu Glu Asn Pro Glu Ile
    210                 215                 220

atc ctc agc gcc cca ccc ctg agt caa ttt gtc atc aca aac acg cac    720
Ile Leu Ser Ala Pro Pro Leu Ser Gln Phe Val Ile Thr Asn Thr His
225                 230                 235                 240

ccc agt ctg cct cag tca gtc agc att att acg cca acc cag ggc gtt    768
Pro Ser Leu Pro Gln Ser Val Ser Ile Ile Thr Pro Thr Gln Gly Val
                245                 250                 255

gtt ccc ggc caa tgt ttt atg gac acg tgg aaa gcg gtg tca cag agc    816
Val Pro Gly Gln Cys Phe Met Asp Thr Trp Lys Ala Val Ser Gln Ser
            260                 265                 270

att cac cac cag gca cag acg cct att ttg gcc gcc gca cta acc ggt    864
Ile His His Gln Ala Gln Thr Pro Ile Leu Ala Ala Ala Leu Thr Gly
        275                 280                 285

tcg aca tct gcg gcc cct ggc ccg cat atc gca tgt tcc cca gtt gcc    912
Ser Thr Ser Ala Ala Pro Gly Pro His Ile Ala Cys Ser Pro Val Ala
    290                 295                 300

ggc acg tct cgg cag gtg gaa ggg tcc gcg ggc gtc gat tgc ggg aaa    960
Gly Thr Ser Arg Gln Val Glu Gly Ser Ala Gly Val Asp Cys Gly Lys
305                 310                 315                 320

cca gca tgc gtt ccg cag ccc gcg tta ccg ccc aat gtc ccc gcc aag   1008
Pro Ala Cys Val Pro Gln Pro Ala Leu Pro Pro Asn Val Pro Ala Lys
                325                 330                 335

agg atg gaa acg gta gca cag ttg gga aac gct ccc gta aaa aac gtc   1056
Arg Met Glu Thr Val Ala Gln Leu Gly Asn Ala Pro Val Lys Asn Val
            340                 345                 350

cac atc gga ggc cgc gta tac gct cca ctg gtt aat ata cca ata ata   1104
His Ile Gly Gly Arg Val Tyr Ala Pro Leu Val Asn Ile Pro Ile Ile
        355                 360                 365

gac tta acg tcc ccg tca ggg tcc ggc cag agt ccg gcc gat atc gcc   1152
Asp Leu Thr Ser Pro Ser Gly Ser Gly Gln Ser Pro Ala Asp Ile Ala
    370                 375                 380
```

aac act cca gag tcc cgc atg gcg gcc ggc tct ccg ccc ttc gcc gaa 1200
Asn Thr Pro Glu Ser Arg Met Ala Ala Gly Ser Pro Pro Phe Ala Glu
385 390 395 400

acc gcc gca acg gtc ccc gct aag aga aag cag cca cgc gag gac gtg 1248
Thr Ala Ala Thr Val Pro Ala Lys Arg Lys Gln Pro Arg Glu Asp Val
405 410 415

gca gac aaa aga ctg aag ggc gac gtt cgg ggc gcc gca aca gta aac 1296
Ala Asp Lys Arg Leu Lys Gly Asp Val Arg Gly Ala Ala Thr Val Asn
420 425 430

cac cct ttc ccg gga ccg tcc ggg atg cgc gtt cgc gag cag ggc tta 1344
His Pro Phe Pro Gly Pro Ser Gly Met Arg Val Arg Glu Gln Gly Leu
435 440 445

ttc gat tta atc gaa agc tcc acg gat gta acc gcg aac gca tct gga 1392
Phe Asp Leu Ile Glu Ser Ser Thr Asp Val Thr Ala Asn Ala Ser Gly
450 455 460

ccc aaa aac gac gac gac atg cta gcg gct atc cta cag gac ctg tat 1440
Pro Lys Asn Asp Asp Asp Met Leu Ala Ala Ile Leu Gln Asp Leu Tyr
465 470 475 480

gga cta cag tcc ccc ccg gcc atc gat tcc ccc tcc agc aac tcg gac 1488
Gly Leu Gln Ser Pro Pro Ala Ile Asp Ser Pro Ser Ser Asn Ser Asp
485 490 495

aat gag gag ata ttt cca gag gtt agt ccg cca tct agc ggc cac gga 1536
Asn Glu Glu Ile Phe Pro Glu Val Ser Pro Pro Ser Ser Gly His Gly
500 505 510

tcg cct tga 1545
Ser Pro
515

<210> 97
<211> 514
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 97
Met Glu Cys Ala Ser Leu Gly Pro Ile Ser Gly Leu Ile Ala Asp Leu
1 5 10 15
Asn Leu Phe Asn Leu Phe Cys Leu Tyr Arg Gly Ser Arg Val Lys Thr
20 25 30
Arg Gly Ala Ala Thr Cys Asn Val Pro Cys Ala Glu Cys Ala Gln Gly
35 40 45
Ile Val Arg Ile Leu Thr Glu Arg Ala Leu Cys Cys Thr Glu Lys Met
50 55 60
Phe Ile Ala Ser Ala Cys Ser Gly Val Val Ile Pro Pro Gln Leu Ala
65 70 75 80
Arg Val Leu His Asp Val Tyr Ala Glu Met Lys Ala Lys Cys Leu Gly
85 90 95
Ala Trp Arg Arg Leu Ile Cys Cys Arg Arg Pro Ile Met Ala Ile Ala
100 105 110
Asp Ser Val Leu Val Thr Tyr Asn Thr Leu Asp Ala Glu Gly Lys Leu
115 120 125
Glu Leu Arg Leu Lys Ala Leu Cys Lys Leu Val Phe Gln Pro Ile Phe

130 135 140
Leu Gln Arg Ile Leu Ala Pro Met Gln Leu Leu Ala Asn Gly Lys Met
145 150 155 160
Val Pro Asp Asn Tyr Phe Thr Ile Thr Gly Thr Ala Glu Lys Arg Arg
165 170 175
Pro Val Val Thr Gly Ser Thr Ser Gly Met Thr Cys Pro Gly Ser Ser
180 185 190
Leu Val Pro Asp Ser Leu Ile Leu Pro Val Cys Glu Pro Gly Leu Leu
195 200 205
Pro Ala Pro Leu Val Asp Leu Ser Asn Val Leu Glu Asn Pro Glu Ile
210 215 220
Ile Leu Ser Ala Pro Pro Leu Ser Gln Phe Val Ile Thr Asn Thr His
225 230 235 240
Pro Ser Leu Pro Gln Ser Val Ser Ile Ile Thr Pro Thr Gln Gly Val
245 250 255
Val Pro Gly Gln Cys Phe Met Asp Thr Trp Lys Ala Val Ser Gln Ser
260 265 270
Ile His His Gln Ala Gln Thr Pro Ile Leu Ala Ala Ala Leu Thr Gly
275 280 285
Ser Thr Ser Ala Ala Pro Gly Pro His Ile Ala Cys Ser Pro Val Ala
290 295 300
Gly Thr Ser Arg Gln Val Glu Gly Ser Ala Gly Val Asp Cys Gly Lys
305 310 315 320
Pro Ala Cys Val Pro Gln Pro Ala Leu Pro Pro Asn Val Pro Ala Lys
325 330 335
Arg Met Glu Thr Val Ala Gln Leu Gly Asn Ala Pro Val Lys Asn Val
340 345 350
His Ile Gly Gly Arg Val Tyr Ala Pro Leu Val Asn Ile Pro Ile Ile
355 360 365
Asp Leu Thr Ser Pro Ser Gly Ser Gly Gln Ser Pro Ala Asp Ile Ala
370 375 380
Asn Thr Pro Glu Ser Arg Met Ala Ala Gly Ser Pro Pro Phe Ala Glu
385 390 395 400
Thr Ala Ala Thr Val Pro Ala Lys Arg Lys Gln Pro Arg Glu Asp Val
405 410 415
Ala Asp Lys Arg Leu Lys Gly Asp Val Arg Gly Ala Ala Thr Val Asn
420 425 430
His Pro Phe Pro Gly Pro Ser Gly Met Arg Val Arg Glu Gln Gly Leu
435 440 445
Phe Asp Leu Ile Glu Ser Ser Thr Asp Val Thr Ala Asn Ala Ser Gly
450 455 460
Pro Lys Asn Asp Asp Asp Met Leu Ala Ala Ile Leu Gln Asp Leu Tyr
465 470 475 480
Gly Leu Gln Ser Pro Pro Ala Ile Asp Ser Pro Ser Ser Asn Ser Asp
485 490 495
Asn Glu Glu Ile Phe Pro Glu Val Ser Pro Pro Ser Ser Gly His Gly
500 505 510
Ser Pro

<210> 98
<211> 534
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(534)

<400> 98

atg cct cgt gtg aaa aca caa ccc aag aga ccc caa gtg ctt gaa ttt 48
Met Pro Arg Val Lys Thr Gln Pro Lys Arg Pro Gln Val Leu Glu Phe
1 5 10 15

atg cca tta gat ctc cac ggt gga aca cac acg gag atg gat tct caa 96
Met Pro Leu Asp Leu His Gly Gly Thr His Thr Glu Met Asp Ser Gln
20 25 30

aac ctg tgt cct gac ggc cag gat ctg ctc ggg tct tat atc tat acg 144
Asn Leu Cys Pro Asp Gly Gln Asp Leu Leu Gly Ser Tyr Ile Tyr Thr
35 40 45

gag aat aac ggc ccg ttt tcc caa ata atg cac aat gga cag agc aat 192
Glu Asn Asn Gly Pro Phe Ser Gln Ile Met His Asn Gly Gln Ser Asn
50 55 60

acc ggg aca ggt gaa agc ttc ggc agc tac gct gcc ggc gac ggt ttt 240
Thr Gly Thr Gly Glu Ser Phe Gly Ser Tyr Ala Ala Gly Asp Gly Phe
65 70 75 80

ctg ggc ggt tct gtg tca ggg atg tat gga aac aac acc gga gag ggc 288
Leu Gly Gly Ser Val Ser Gly Met Tyr Gly Asn Asn Thr Gly Glu Gly
85 90 95

gcg tgt tct aaa aga ccg tcc gcg tgc cgt aaa cgc tcg gct gca cta 336
Ala Cys Ser Lys Arg Pro Ser Ala Cys Arg Lys Arg Ser Ala Ala Leu
100 105 110

att cac gcg gcg tcc gag gcg tct gtg gcc gag caa ggc acc tca cag 384
Ile His Ala Ala Ser Glu Ala Ser Val Ala Glu Gln Gly Thr Ser Gln
115 120 125

ggg gca cat gcc gta tct gac cgg ata ggc aga gac ggt ggc gct gac 432
Gly Ala His Ala Val Ser Asp Arg Ile Gly Arg Asp Gly Gly Ala Asp
130 135 140

aat aga cta ctc aag gtg agt gcg cgg ctg tcg gac aaa aca aag agc 480
Asn Arg Leu Leu Lys Val Ser Ala Arg Leu Ser Asp Lys Thr Lys Ser
145 150 155 160

gcc ctt cgc agc cat cct tgc ttg cgt tgc tat tct ttg atg ttt aac 528
Ala Leu Arg Ser His Pro Cys Leu Arg Cys Tyr Ser Leu Met Phe Asn
165 170 175

acg taa 534
Thr

<210> 99
<211> 177
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 99
Met Pro Arg Val Lys Thr Gln Pro Lys Arg Pro Gln Val Leu Glu Phe
1 5 10 15
Met Pro Leu Asp Leu His Gly Gly Thr His Thr Glu Met Asp Ser Gln
20 25 30
Asn Leu Cys Pro Asp Gly Gln Asp Leu Leu Gly Ser Tyr Ile Tyr Thr
35 40 45
Glu Asn Asn Gly Pro Phe Ser Gln Ile Met His Asn Gly Gln Ser Asn

50 55 60
Thr Gly Thr Gly Glu Ser Phe Gly Ser Tyr Ala Ala Gly Asp Gly Phe
65 70 75 80
Leu Gly Gly Ser Val Ser Gly Met Tyr Gly Asn Asn Thr Gly Glu Gly
85 90 95
Ala Cys Ser Lys Arg Pro Ser Ala Cys Arg Lys Arg Ser Ala Ala Leu
100 105 110
Ile His Ala Ala Ser Glu Ala Ser Val Ala Glu Gln Gly Thr Ser Gln
115 120 125
Gly Ala His Ala Val Ser Asp Arg Ile Gly Arg Asp Gly Gly Ala Asp
130 135 140
Asn Arg Leu Leu Lys Val Ser Ala Arg Leu Ser Asp Lys Thr Lys Ser
145 150 155 160
Ala Leu Arg Ser His Pro Cys Leu Arg Cys Tyr Ser Leu Met Phe Asn
165 170 175
Thr

<210> 100
<211> 693
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(693)

<400> 100
atg gga ttt ggg aac ata cgt ctg gga tgg agg tta tgc ttc atg gtc 48
Met Gly Phe Gly Asn Ile Arg Leu Gly Trp Arg Leu Cys Phe Met Val
1 5 10 15

tgg gtg gcg tgg att gca cgg gga cgg tcg gtg tgc cca acc tgg cac 96
Trp Val Ala Trp Ile Ala Arg Gly Arg Ser Val Cys Pro Thr Trp His
20 25 30

ctg aca gat ggg aaa tac gag gcg gta tac agg cac tac ctc gaa gag 144
Leu Thr Asp Gly Lys Tyr Glu Ala Val Tyr Arg His Tyr Leu Glu Glu
35 40 45

tgc cgc aaa cat gaa ggc tcg ggg agc ctg gac ggt tcc gga cag aca 192
Cys Arg Lys His Glu Gly Ser Gly Ser Leu Asp Gly Ser Gly Gln Thr
50 55 60

aag ggg tct gga acc aaa gca acc acc gaa gct aat ata tcg ata aga 240
Lys Gly Ser Gly Thr Lys Ala Thr Thr Glu Ala Asn Ile Ser Ile Arg
65 70 75 80

cct aac gtt gtc aca tca ggt caa aat aaa gag ccg cct ggg aca gca 288
Pro Asn Val Val Thr Ser Gly Gln Asn Lys Glu Pro Pro Gly Thr Ala
85 90 95

ccg agg gcc gaa tca tca cac gac ctg cca cgc atc aag cag gtt aac 336
Pro Arg Ala Glu Ser Ser His Asp Leu Pro Arg Ile Lys Gln Val Asn
100 105 110

gct ctc cga tta tca acc ccg gaa ttg gcg caa cca ctc ccg gta gta 384
Ala Leu Arg Leu Ser Thr Pro Glu Leu Ala Gln Pro Leu Pro Val Val
115 120 125

```
aaa tcg act ccg cgc gag tca cag tca ggt ggg aca ccc tgg aac gcg      432
Lys Ser Thr Pro Arg Glu Ser Gln Ser Gly Gly Thr Pro Trp Asn Ala
    130                 135                 140

cgc ccc cac gcg ttc att atg cac aca aac gac atg ctc aac cca tct      480
Arg Pro His Ala Phe Ile Met His Thr Asn Asp Met Leu Asn Pro Ser
145                 150                 155                 160

gtg gtc ctg tct ttc aga gcc atc cgt gcg cgg tcc aca cgc gat acc      528
Val Val Leu Ser Phe Arg Ala Ile Arg Ala Arg Ser Thr Arg Asp Thr
                165                 170                 175

gag cag tcc gtt cgc gat cgg aac acg gtc acg acc agc tat cgt acc      576
Glu Gln Ser Val Arg Asp Arg Asn Thr Val Thr Thr Ser Tyr Arg Thr
            180                 185                 190

cct ggc cgc cct tcc ctc ttt caa gcc aga ccc tcg tct cac ggt gcg      624
Pro Gly Arg Pro Ser Leu Phe Gln Ala Arg Pro Ser Ser His Gly Ala
        195                 200                 205

cgt cta ccc cct tcg ccc cga acg atg gca aga tac gcc gag tcg cga      672
Arg Leu Pro Pro Ser Pro Arg Thr Met Ala Arg Tyr Ala Glu Ser Arg
    210                 215                 220

aca ata tgc gac caa aat tga                                          693
Thr Ile Cys Asp Gln Asn
225                 230
```

<210> 101
<211> 230
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 101

```
Met Gly Phe Gly Asn Ile Arg Leu Gly Trp Arg Leu Cys Phe Met Val
1               5                   10                  15
Trp Val Ala Trp Ile Ala Arg Gly Arg Ser Val Cys Pro Thr Trp His
            20                  25                  30
Leu Thr Asp Gly Lys Tyr Glu Ala Val Tyr Arg His Tyr Leu Glu Glu
        35                  40                  45
Cys Arg Lys His Glu Gly Ser Gly Ser Leu Asp Gly Ser Gly Gln Thr
    50                  55                  60
Lys Gly Ser Gly Thr Lys Ala Thr Thr Glu Ala Asn Ile Ser Ile Arg
65                  70                  75                  80
Pro Asn Val Val Thr Ser Gly Gln Asn Lys Glu Pro Pro Gly Thr Ala
                85                  90                  95
Pro Arg Ala Glu Ser Ser His Asp Leu Pro Arg Ile Lys Gln Val Asn
            100                 105                 110
Ala Leu Arg Leu Ser Thr Pro Glu Leu Ala Gln Pro Leu Pro Val Val
        115                 120                 125
Lys Ser Thr Pro Arg Glu Ser Gln Ser Gly Gly Thr Pro Trp Asn Ala
    130                 135                 140
Arg Pro His Ala Phe Ile Met His Thr Asn Asp Met Leu Asn Pro Ser
145                 150                 155                 160
Val Val Leu Ser Phe Arg Ala Ile Arg Ala Arg Ser Thr Arg Asp Thr
                165                 170                 175
Glu Gln Ser Val Arg Asp Arg Asn Thr Val Thr Thr Ser Tyr Arg Thr
            180                 185                 190
Pro Gly Arg Pro Ser Leu Phe Gln Ala Arg Pro Ser Ser His Gly Ala
        195                 200                 205
```

Arg Leu Pro Pro Ser Pro Arg Thr Met Ala Arg Tyr Ala Glu Ser Arg
210 215 220

Thr Ile Cys Asp Gln Asn
225 230

<210> 102
<211> 420
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(420)

<400> 102
atg tct tcc acg cgt ccc aag acg cgc gcc ccc aaa aaa gag ctg aca 48
Met Ser Ser Thr Arg Pro Lys Thr Arg Ala Pro Lys Lys Glu Leu Thr
1 5 10 15

atg gag gag tta gcc gct cag gtg caa aaa ttg tcc gtt gaa aac aag 96
Met Glu Glu Leu Ala Ala Gln Val Gln Lys Leu Ser Val Glu Asn Lys
20 25 30

cag ctc aaa aag ctg ata aat tct ggg gat ccc acc agg tcc ggc tcc 144
Gln Leu Lys Lys Leu Ile Asn Ser Gly Asp Pro Thr Arg Ser Gly Ser
35 40 45

gac ccc gtc atc tct aac acc gaa aag gag gcc aaa atc gca gct gcc 192
Asp Pro Val Ile Ser Asn Thr Glu Lys Glu Ala Lys Ile Ala Ala Ala
50 55 60

gtt tcg gcc tta tgt aac gtg gca acc cgt aag att gaa gct aag gtg 240
Val Ser Ala Leu Cys Asn Val Ala Thr Arg Lys Ile Glu Ala Lys Val
65 70 75 80

agg gcg gcc acg gcc aag gcc gtg acg cgc ggg cag atg gaa gac gcc 288
Arg Ala Ala Thr Ala Lys Ala Val Thr Arg Gly Gln Met Glu Asp Ala
85 90 95

ctc gcc gga att agt atc cgc gtg gac gtg tca atg gac gag acc acc 336
Leu Ala Gly Ile Ser Ile Arg Val Asp Val Ser Met Asp Glu Thr Thr
100 105 110

aga ggc gga atc gct gcc agc gcg gac ggg gcg ctc agg agg agg cgc 384
Arg Gly Gly Ile Ala Ala Ser Ala Asp Gly Ala Leu Arg Arg Arg Arg
115 120 125

gcc cag tcc agg acc aga aat aac gac gcg gac tag 420
Ala Gln Ser Arg Thr Arg Asn Asn Asp Ala Asp
130 135 140

<210> 103
<211> 139
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 103
Met Ser Ser Thr Arg Pro Lys Thr Arg Ala Pro Lys Lys Glu Leu Thr

```
1               5                   10                  15
Met Glu Glu Leu Ala Ala Gln Val Gln Lys Leu Ser Val Glu Asn Lys
            20                  25                  30
Gln Leu Lys Lys Leu Ile Asn Ser Gly Asp Pro Thr Arg Ser Gly Ser
        35                  40                  45
Asp Pro Val Ile Ser Asn Thr Glu Lys Glu Ala Lys Ile Ala Ala Ala
    50                  55                  60
Val Ser Ala Leu Cys Asn Val Ala Thr Arg Lys Ile Glu Ala Lys Val
65                  70                  75                  80
Arg Ala Ala Thr Ala Lys Ala Val Thr Arg Gly Gln Met Glu Asp Ala
                85                  90                  95
Leu Ala Gly Ile Ser Ile Arg Val Asp Val Ser Met Asp Glu Thr Thr
            100                 105                 110
Arg Gly Gly Ile Ala Ala Ser Ala Asp Gly Ala Leu Arg Arg Arg Arg
        115                 120                 125
Ala Gln Ser Arg Thr Arg Asn Asn Asp Ala Asp
    130                 135
```

<210> 104
<211> 315
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(315)

<400> 104

```
atg aca ggg tca att gtg ctc gcg tta gcc ctg ttg gcg tgc tta tat      48
Met Thr Gly Ser Ile Val Leu Ala Leu Ala Leu Leu Ala Cys Leu Tyr
1               5                   10                  15

ctg tgt ctg ccg gtg tgc gcc acc gtc acc acc tcc tcg act aca gga      96
Leu Cys Leu Pro Val Cys Ala Thr Val Thr Thr Ser Ser Thr Thr Gly
            20                  25                  30

acg ggc acg cct cct gtg act aca aac ccc tcc gcg gct ccc tcc gta     144
Thr Gly Thr Pro Pro Val Thr Thr Asn Pro Ser Ala Ala Pro Ser Val
        35                  40                  45

acg ccc tct ttt tat gac tat gat tgc tcg gcg gac acg tac cag ccc     192
Thr Pro Ser Phe Tyr Asp Tyr Asp Cys Ser Ala Asp Thr Tyr Gln Pro
    50                  55                  60

gtg ctg tct agt ttt tcg agt ata tgg gcg gtg atc aac tct gtt ctg     240
Val Leu Ser Ser Phe Ser Ser Ile Trp Ala Val Ile Asn Ser Val Leu
65                  70                  75                  80

gtc gcg gtt gca aca ttt ctg tac cta aca tat atg tgc ttt ttt aaa     288
Val Ala Val Ala Thr Phe Leu Tyr Leu Thr Tyr Met Cys Phe Phe Lys
                85                  90                  95

ttt gtt gaa acc gtg gcg cac gag tga                                 315
Phe Val Glu Thr Val Ala His Glu
            100                 105
```

<210> 105
<211> 104

```
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 105
Met Thr Gly Ser Ile Val Leu Ala Leu Ala Leu Leu Ala Cys Leu Tyr
1               5                   10                  15
Leu Cys Leu Pro Val Cys Ala Thr Val Thr Thr Ser Ser Thr Thr Gly
            20                  25                  30
Thr Gly Thr Pro Pro Val Thr Thr Asn Pro Ser Ala Ala Pro Ser Val
        35                  40                  45
Thr Pro Ser Phe Tyr Asp Tyr Asp Cys Ser Ala Asp Thr Tyr Gln Pro
    50                  55                  60
Val Leu Ser Ser Phe Ser Ser Ile Trp Ala Val Ile Asn Ser Val Leu
65                  70                  75                  80
Val Ala Val Ala Thr Phe Leu Tyr Leu Thr Tyr Met Cys Phe Phe Lys
                85                  90                  95
Phe Val Glu Thr Val Ala His Glu
            100


<210> 106
<211> 873
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(873)

<400> 106
atg gca gaa gtg acc gcc cat acg gtt ccg tac gcc ttt gac tcc tgt      48
Met Ala Glu Val Thr Ala His Thr Val Pro Tyr Ala Phe Asp Ser Cys
1               5                   10                  15

aag ttt gaa ata ata cca aaa aac aac tca tct cgt ata gcg tta cgc      96
Lys Phe Glu Ile Ile Pro Lys Asn Asn Ser Ser Arg Ile Ala Leu Arg
            20                  25                  30

aac aaa ttt ccc gtc gtg gtc aaa ccg gga gaa cct cta gtc gtg ccc     144
Asn Lys Phe Pro Val Val Val Lys Pro Gly Glu Pro Leu Val Val Pro
        35                  40                  45

ctg gga cta aaa att atc cgc gcg ccc cag tgc gca ttc ttt ctc agc     192
Leu Gly Leu Lys Ile Ile Arg Ala Pro Gln Cys Ala Phe Phe Leu Ser
    50                  55                  60

gga gcg ccg acg gat gag gtg tat tac cac acc gga ctt ata gat caa     240
Gly Ala Pro Thr Asp Glu Val Tyr Tyr His Thr Gly Leu Ile Asp Gln
65                  70                  75                  80

ggg tat cgt gga gaa ata aaa ctc atc gtt ctc aat aag act aaa cag     288
Gly Tyr Arg Gly Glu Ile Lys Leu Ile Val Leu Asn Lys Thr Lys Gln
                85                  90                  95

gtc gtc aca ctt tac cgg gga gaa gtc aac gtc tca ctc atc gcg ttt     336
Val Val Thr Leu Tyr Arg Gly Glu Val Asn Val Ser Leu Ile Ala Phe
            100                 105                 110

atg tac gcc tcg ccc ggg ccc cta aaa tgc ccg ata ctg aac cta cca     384
Met Tyr Ala Ser Pro Gly Pro Leu Lys Cys Pro Ile Leu Asn Leu Pro
```

115 120 125

cat tac tct ctg gac gct gga ttt gat gtt acg tcg cca cat gcc atg 432
His Tyr Ser Leu Asp Ala Gly Phe Asp Val Thr Ser Pro His Ala Met
130 135 140

acc ata ccc cca acc gac aga acc ccg ttt acg ttg agt ttg tat tat 480
Thr Ile Pro Pro Thr Asp Arg Thr Pro Phe Thr Leu Ser Leu Tyr Tyr
145 150 155 160

aag agt cca cag ctc agc acc ccc cac gtc cca cta atc gtg ggg cgc 528
Lys Ser Pro Gln Leu Ser Thr Pro His Val Pro Leu Ile Val Gly Arg
165 170 175

tct gga ctg gcg aca aag ggg ctt acc gtc gac gcg acc aaa tgg acc 576
Ser Gly Leu Ala Thr Lys Gly Leu Thr Val Asp Ala Thr Lys Trp Thr
180 185 190

cag tcg ttg gta cat ctg agg ttt tat aac ttt acc aaa gag cca ata 624
Gln Ser Leu Val His Leu Arg Phe Tyr Asn Phe Thr Lys Glu Pro Ile
195 200 205

gat ata cca gca aac agc cgc ata tgt cag gta gtg ttt atc cac gaa 672
Asp Ile Pro Ala Asn Ser Arg Ile Cys Gln Val Val Phe Ile His Glu
210 215 220

gac cac gtc cca agc ggg tgg aac att ctt aga tcc cgc gta caa ctc 720
Asp His Val Pro Ser Gly Trp Asn Ile Leu Arg Ser Arg Val Gln Leu
225 230 235 240

ggc agc acc ctc cag ata tcg tgg gcc aaa atc agg ttt acc gac gtg 768
Gly Ser Thr Leu Gln Ile Ser Trp Ala Lys Ile Arg Phe Thr Asp Val
245 250 255

gcc acg ctc ccc aaa acc cac ccg ctc aac tcc cgc cac act caa agc 816
Ala Thr Leu Pro Lys Thr His Pro Leu Asn Ser Arg His Thr Gln Ser
260 265 270

caa acc gaa ccg gag acc gcc cgc ggc gca aag ggg ttg ggg tcg tca 864
Gln Thr Glu Pro Glu Thr Ala Arg Gly Ala Lys Gly Leu Gly Ser Ser
275 280 285

ggg tta taa 873
Gly Leu
290

<210> 107
<211> 290
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 107
Met Ala Glu Val Thr Ala His Thr Val Pro Tyr Ala Phe Asp Ser Cys
1 5 10 15
Lys Phe Glu Ile Ile Pro Lys Asn Asn Ser Ser Arg Ile Ala Leu Arg
20 25 30
Asn Lys Phe Pro Val Val Val Lys Pro Gly Glu Pro Leu Val Val Pro
35 40 45
Leu Gly Leu Lys Ile Ile Arg Ala Pro Gln Cys Ala Phe Phe Leu Ser
50 55 60

Gly Ala Pro Thr Asp Glu Val Tyr Tyr His Thr Gly Leu Ile Asp Gln
65 70 75 80
Gly Tyr Arg Gly Glu Ile Lys Leu Ile Val Leu Asn Lys Thr Lys Gln
85 90 95
Val Val Thr Leu Tyr Arg Gly Glu Val Asn Val Ser Leu Ile Ala Phe
100 105 110
Met Tyr Ala Ser Pro Gly Pro Leu Lys Cys Pro Ile Leu Asn Leu Pro
115 120 125
His Tyr Ser Leu Asp Ala Gly Phe Asp Val Thr Ser Pro His Ala Met
130 135 140
Thr Ile Pro Pro Thr Asp Arg Thr Pro Phe Thr Leu Ser Leu Tyr Tyr
145 150 155 160
Lys Ser Pro Gln Leu Ser Thr Pro His Val Pro Leu Ile Val Gly Arg
165 170 175
Ser Gly Leu Ala Thr Lys Gly Leu Thr Val Asp Ala Thr Lys Trp Thr
180 185 190
Gln Ser Leu Val His Leu Arg Phe Tyr Asn Phe Thr Lys Glu Pro Ile
195 200 205
Asp Ile Pro Ala Asn Ser Arg Ile Cys Gln Val Val Phe Ile His Glu
210 215 220
Asp His Val Pro Ser Gly Trp Asn Ile Leu Arg Ser Arg Val Gln Leu
225 230 235 240
Gly Ser Thr Leu Gln Ile Ser Trp Ala Lys Ile Arg Phe Thr Asp Val
245 250 255
Ala Thr Leu Pro Lys Thr His Pro Leu Asn Ser Arg His Thr Gln Ser
260 265 270
Gln Thr Glu Pro Glu Thr Ala Arg Gly Ala Lys Gly Leu Gly Ser Ser
275 280 285
Gly Leu
290

<210> 108
<211> 633
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(633)

<400> 108
atg gcc gcg cct ggt tcg ttt tgg acg tgt tgc ggg ttt agt ccg ttt 48
Met Ala Ala Pro Gly Ser Phe Trp Thr Cys Cys Gly Phe Ser Pro Phe
1 5 10 15

gga cgc gtg gga tgt caa tat cgt ccc ctg cca gat cct ctt aac gag 96
Gly Arg Val Gly Cys Gln Tyr Arg Pro Leu Pro Asp Pro Leu Asn Glu
20 25 30

tgt cct act cac tgg cgg acc gaa ata gcc atg ggt ctg ccc ccg gga 144
Cys Pro Thr His Trp Arg Thr Glu Ile Ala Met Gly Leu Pro Pro Gly
35 40 45

gtg gat atg gga gat gtg aag cag gca gag atg tgc acc gcc gcc ctt 192
Val Asp Met Gly Asp Val Lys Gln Ala Glu Met Cys Thr Ala Ala Leu
50 55 60

cgg cag acg tat ctg ttg gcc gtt caa tcc aat aaa ata acc gag tat 240
Arg Gln Thr Tyr Leu Leu Ala Val Gln Ser Asn Lys Ile Thr Glu Tyr

65 70 75 80

ctt cgt cgt ttt gac gcg gct agg gta ccc gcg ggg tgc cag gag acg 288
Leu Arg Arg Phe Asp Ala Ala Arg Val Pro Ala Gly Cys Gln Glu Thr
85 90 95

gtt agg att caa ata tct aag cta aag tct ata cag aac gtt atc tgg 336
Val Arg Ile Gln Ile Ser Lys Leu Lys Ser Ile Gln Asn Val Ile Trp
100 105 110

aac gcc atg ttg tcc ctg gct ata ggg gat att acg gtg gac gag agc 384
Asn Ala Met Leu Ser Leu Ala Ile Gly Asp Ile Thr Val Asp Glu Ser
115 120 125

gcc ttt cac gcg ctt ttg aat aaa cgg gcc gac gag acc gtg tca ctc 432
Ala Phe His Ala Leu Leu Asn Lys Arg Ala Asp Glu Thr Val Ser Leu
130 135 140

ctg gaa atg gaa aaa ctg gct acg acg atc gcg tcc gat gat tct gtc 480
Leu Glu Met Glu Lys Leu Ala Thr Thr Ile Ala Ser Asp Asp Ser Val
145 150 155 160

acg tgg gct gcg gaa att aat aat gtg ctg gta gac acc gaa gcg tcc 528
Thr Trp Ala Ala Glu Ile Asn Asn Val Leu Val Asp Thr Glu Ala Ser
165 170 175

tcc aac ccc tcc cat ccg gtg att cga caa ccg acg ccc cag ttg gcg 576
Ser Asn Pro Ser His Pro Val Ile Arg Gln Pro Thr Pro Gln Leu Ala
180 185 190

gtg gcg gat aat atc gtg cca gac aaa ata ata cag gat gcg caa gct 624
Val Ala Asp Asn Ile Val Pro Asp Lys Ile Ile Gln Asp Ala Gln Ala
195 200 205

gat ggt taa 633
Asp Gly
210

<210> 109
<211> 210
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 109
Met Ala Ala Pro Gly Ser Phe Trp Thr Cys Cys Gly Phe Ser Pro Phe
1 5 10 15
Gly Arg Val Gly Cys Gln Tyr Arg Pro Leu Pro Asp Pro Leu Asn Glu
20 25 30
Cys Pro Thr His Trp Arg Thr Glu Ile Ala Met Gly Leu Pro Pro Gly
35 40 45
Val Asp Met Gly Asp Val Lys Gln Ala Glu Met Cys Thr Ala Ala Leu
50 55 60
Arg Gln Thr Tyr Leu Leu Ala Val Gln Ser Asn Lys Ile Thr Glu Tyr
65 70 75 80
Leu Arg Arg Phe Asp Ala Ala Arg Val Pro Ala Gly Cys Gln Glu Thr
85 90 95
Val Arg Ile Gln Ile Ser Lys Leu Lys Ser Ile Gln Asn Val Ile Trp
100 105 110
Asn Ala Met Leu Ser Leu Ala Ile Gly Asp Ile Thr Val Asp Glu Ser
115 120 125

Ala Phe His Ala Leu Leu Asn Lys Arg Ala Asp Glu Thr Val Ser Leu
130 135 140
Leu Glu Met Glu Lys Leu Ala Thr Thr Ile Ala Ser Asp Asp Ser Val
145 150 155 160
Thr Trp Ala Ala Glu Ile Asn Asn Val Leu Val Asp Thr Glu Ala Ser
165 170 175
Ser Asn Pro Ser His Pro Val Ile Arg Gln Pro Thr Pro Gln Leu Ala
180 185 190
Val Ala Asp Asn Ile Val Pro Asp Lys Ile Ile Gln Asp Ala Gln Ala
195 200 205
Asp Gly
210

<210> 110
<211> 2487
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2487)

<400> 110
atg gta gat gaa att agg gca att ttc tct act agt gga gat atg gcc 48
Met Val Asp Glu Ile Arg Ala Ile Phe Ser Thr Ser Gly Asp Met Ala
1 5 10 15

gaa gta att acg gat ata ctg act gaa acg caa gca acg gcg tcc ttc 96
Glu Val Ile Thr Asp Ile Leu Thr Glu Thr Gln Ala Thr Ala Ser Phe
20 25 30

ttc tgc gtg ctc cac gat cgg ggc gac gcg cct ata aat act cca cat 144
Phe Cys Val Leu His Asp Arg Gly Asp Ala Pro Ile Asn Thr Pro His
35 40 45

gcc gta att aaa ctc tgc ctg ccc gcc aag cgc cca ggc ggc ggg cca 192
Ala Val Ile Lys Leu Cys Leu Pro Ala Lys Arg Pro Gly Gly Gly Pro
50 55 60

agg tgt tta ccg ttg atg gtg ctg aac cta ccg gcg tgg cag gtt aat 240
Arg Cys Leu Pro Leu Met Val Leu Asn Leu Pro Ala Trp Gln Val Asn
65 70 75 80

cta ttc cta aca ggt gac gca cca ttg acc tcg gat aac att aaa gac 288
Leu Phe Leu Thr Gly Asp Ala Pro Leu Thr Ser Asp Asn Ile Lys Asp
85 90 95

cgc att gac ctg gct cag acc gag gaa ata ctc gaa ccc ata tta agc 336
Arg Ile Asp Leu Ala Gln Thr Glu Glu Ile Leu Glu Pro Ile Leu Ser
100 105 110

gta ctg gca tgc aaa cgg tcc gcg cag cag acc aaa cat gac tcg ttt 384
Val Leu Ala Cys Lys Arg Ser Ala Gln Gln Thr Lys His Asp Ser Phe
115 120 125

aaa tct aag gtg gcc tgg ttc agg gca aag ttc gtc tcg gcc cta aga 432
Lys Ser Lys Val Ala Trp Phe Arg Ala Lys Phe Val Ser Ala Leu Arg
130 135 140

aaa gtg tac aaa atg acc cca tcc ccc tat tgg atg ata acg ctg ctg 480
Lys Val Tyr Lys Met Thr Pro Ser Pro Tyr Trp Met Ile Thr Leu Leu
145 150 155 160

ggc tcg ttc gag gca tcg ttt gta ctg gcc ggc acg ttt tat ttt ttc 528
Gly Ser Phe Glu Ala Ser Phe Val Leu Ala Gly Thr Phe Tyr Phe Phe
165 170 175

cag tct cat ata tgc acg gcg gag acg ctg gtg cat cta acg agg tta 576
Gln Ser His Ile Cys Thr Ala Glu Thr Leu Val His Leu Thr Arg Leu
180 185 190

ttc agc tcc agt cag ggc cag agc ttg gtc acc gta aac acc tat gac 624
Phe Ser Ser Ser Gln Gly Gln Ser Leu Val Thr Val Asn Thr Tyr Asp
195 200 205

gaa tta ggg cgc gta ttc ggg cgg tca gat ttc ctt gga att gtt cca 672
Glu Leu Gly Arg Val Phe Gly Arg Ser Asp Phe Leu Gly Ile Val Pro
210 215 220

aat ttt tgg gcg tat cta aaa tac aaa atg caa cag gac gac gtg gag 720
Asn Phe Trp Ala Tyr Leu Lys Tyr Lys Met Gln Gln Asp Asp Val Glu
225 230 235 240

tcc agg gcc atc gat caa aca atc aac tcc ata cgt ggg ggg ttg atg 768
Ser Arg Ala Ile Asp Gln Thr Ile Asn Ser Ile Arg Gly Gly Leu Met
245 250 255

ctg tca cct cag gac ctc gta cac ttc atc tac ctg tca ttt tac gag 816
Leu Ser Pro Gln Asp Leu Val His Phe Ile Tyr Leu Ser Phe Tyr Glu
260 265 270

tgc atg aac gca cag acg ttc ctg tcg tac tct cgt acc acg tcc agt 864
Cys Met Asn Ala Gln Thr Phe Leu Ser Tyr Ser Arg Thr Thr Ser Ser
275 280 285

tta ccg acc ccc gca acg gtt aac cct ccg cag ctg tgt cgg cgc tta 912
Leu Pro Thr Pro Ala Thr Val Asn Pro Pro Gln Leu Cys Arg Arg Leu
290 295 300

gaa gcg gat ttt aag gag cac gtg atg gca tat tac aat aaa gca agc 960
Glu Ala Asp Phe Lys Glu His Val Met Ala Tyr Tyr Asn Lys Ala Ser
305 310 315 320

tac ctg agt act tac ata acc att tta acc gta ccc gct cct ctc ccg 1008
Tyr Leu Ser Thr Tyr Ile Thr Ile Leu Thr Val Pro Ala Pro Leu Pro
325 330 335

gac gga tac gaa aac ttt cag gaa cta gca tgt caa tac tgg tgt gga 1056
Asp Gly Tyr Glu Asn Phe Gln Glu Leu Ala Cys Gln Tyr Trp Cys Gly
340 345 350

caa tcg aga gac gtg gcg gaa atc atg act aga att aat gac cag tat 1104
Gln Ser Arg Asp Val Ala Glu Ile Met Thr Arg Ile Asn Asp Gln Tyr
355 360 365

cca cag ctg aat cta aca aaa gat tta tcc ggt ctc ctt gac cta gcg 1152
Pro Gln Leu Asn Leu Thr Lys Asp Leu Ser Gly Leu Leu Asp Leu Ala
370 375 380

gcc cta gat cag tat tcc gga ggc ccc aag gaa aac ctt ttt acg gtg 1200

Ala Leu Asp Gln Tyr Ser Gly Gly Pro Lys Glu Asn Leu Phe Thr Val
385 390 395 400

gct tcg aga att ccc aca tac agg tgc gag ttt ttg aat aaa cag tat 1248
Ala Ser Arg Ile Pro Thr Tyr Arg Cys Glu Phe Leu Asn Lys Gln Tyr
405 410 415

ttt gtt cta atg cac gca gac tgt ata gac gcg tac tgg aaa caa aac 1296
Phe Val Leu Met His Ala Asp Cys Ile Asp Ala Tyr Trp Lys Gln Asn
420 425 430

att atc gtg ccc gaa gac gcc caa ttg caa ggc ctg acg gat caa gat 1344
Ile Ile Val Pro Glu Asp Ala Gln Leu Gln Gly Leu Thr Asp Gln Asp
435 440 445

ctc acc tcc agg ata ttt tac tgc gac ctc ggc cta tct cta cct aca 1392
Leu Thr Ser Arg Ile Phe Tyr Cys Asp Leu Gly Leu Ser Leu Pro Thr
450 455 460

ttt aaa caa caa att cta gtt tca cgc cac gag tat ttc aac cca cga 1440
Phe Lys Gln Gln Ile Leu Val Ser Arg His Glu Tyr Phe Asn Pro Arg
465 470 475 480

ctc cca gtt tat aga tgg gtg ttg gat ttt gat cta aag gta acc gaa 1488
Leu Pro Val Tyr Arg Trp Val Leu Asp Phe Asp Leu Lys Val Thr Glu
485 490 495

ggt aga cgg act cta aac gat atc tac aac ata tgc gta acg ctg cgg 1536
Gly Arg Arg Thr Leu Asn Asp Ile Tyr Asn Ile Cys Val Thr Leu Arg
500 505 510

cag gtc ata tta gaa acg ttg cag tta ata ggt cca cta aaa ccg aac 1584
Gln Val Ile Leu Glu Thr Leu Gln Leu Ile Gly Pro Leu Lys Pro Asn
515 520 525

cac ccc gtg tat ttt ttt aag tcg gcc tgt cca gcc gta acc tgg ccg 1632
His Pro Val Tyr Phe Phe Lys Ser Ala Cys Pro Ala Val Thr Trp Pro
530 535 540

gat gat atc tca gac acc gcg ttc tgt cac tgt gac gca aaa ata ggg 1680
Asp Asp Ile Ser Asp Thr Ala Phe Cys His Cys Asp Ala Lys Ile Gly
545 550 555 560

atg aga att gta acc ccg ttt ccg agc gga tac tgt ttg gtg ggg tcc 1728
Met Arg Ile Val Thr Pro Phe Pro Ser Gly Tyr Cys Leu Val Gly Ser
565 570 575

gct cca ctc gtg tcc ctg acc gac att tta aac cgc gtg gtt aaa ctg 1776
Ala Pro Leu Val Ser Leu Thr Asp Ile Leu Asn Arg Val Val Lys Leu
580 585 590

gac acg cgg cta gcc tca gag tat cct ggg atc tta gaa gat aaa gga 1824
Asp Thr Arg Leu Ala Ser Glu Tyr Pro Gly Ile Leu Glu Asp Lys Gly
595 600 605

cca ttt gac tct ggc att tac gcc aag ggg cga tgc gtg cgc gta cct 1872
Pro Phe Asp Ser Gly Ile Tyr Ala Lys Gly Arg Cys Val Arg Val Pro
610 615 620

cac tgc tat aag gtt ggc ccg ggg ggc gag ctg tca cgg ctc cta aaa 1920
His Cys Tyr Lys Val Gly Pro Gly Gly Glu Leu Ser Arg Leu Leu Lys

625 630 635 640

att att atc tgt cac ccc gaa gag tca gat aaa tcg gcg tat ttg aaa 1968
Ile Ile Ile Cys His Pro Glu Glu Ser Asp Lys Ser Ala Tyr Leu Lys
645 650 655

aac gcg ttt aag gtt tct aat ctg cta cat cac gcc cct gga gac tct 2016
Asn Ala Phe Lys Val Ser Asn Leu Leu His His Ala Pro Gly Asp Ser
660 665 670

gtc acc aaa aac ggc cac ctg gtg tac gcg atc act gac gaa aac gag 2064
Val Thr Lys Asn Gly His Leu Val Tyr Ala Ile Thr Asp Glu Asn Glu
675 680 685

ggc ttt cta gaa agc aaa acc aag aat aat ctt ccc aaa aca atc acg 2112
Gly Phe Leu Glu Ser Lys Thr Lys Asn Asn Leu Pro Lys Thr Ile Thr
690 695 700

gac ctg gcc gaa aaa att gag cga acc aca gaa aaa cca cta att gat 2160
Asp Leu Ala Glu Lys Ile Glu Arg Thr Thr Glu Lys Pro Leu Ile Asp
705 710 715 720

tgg gca gcg acc gca gtg tgg cca aaa cta cac gat acc ata cag cgt 2208
Trp Ala Ala Thr Ala Val Trp Pro Lys Leu His Asp Thr Ile Gln Arg
725 730 735

ttc ttt ccg gat gac cgc atc ggc caa ttt gca tct gtg agc ttc atg 2256
Phe Phe Pro Asp Asp Arg Ile Gly Gln Phe Ala Ser Val Ser Phe Met
740 745 750

cac tcc gga gac aac att ata caa gtt aaa cca caa aaa gga aac aac 2304
His Ser Gly Asp Asn Ile Ile Gln Val Lys Pro Gln Lys Gly Asn Asn
755 760 765

ttt ttc tgt att aat cat aag cac cga aac cac act cag aca gtc cgt 2352
Phe Phe Cys Ile Asn His Lys His Arg Asn His Thr Gln Thr Val Arg
770 775 780

gta ttt tta acc cta cat tcc aca aaa gaa agc gag gtc acg gtg acc 2400
Val Phe Leu Thr Leu His Ser Thr Lys Glu Ser Glu Val Thr Val Thr
785 790 795 800

ttt atg agt cag tgc ttt gct gca aaa tgt aat cat aat agc cca act 2448
Phe Met Ser Gln Cys Phe Ala Ala Lys Cys Asn His Asn Ser Pro Thr
805 810 815

gca cat ttt tca ttt atg gta ccc ata acc ggc acg taa 2487
Ala His Phe Ser Phe Met Val Pro Ile Thr Gly Thr
820 825

<210> 111
<211> 828
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 111
Met Val Asp Glu Ile Arg Ala Ile Phe Ser Thr Ser Gly Asp Met Ala
1 5 10 15
Glu Val Ile Thr Asp Ile Leu Thr Glu Thr Gln Ala Thr Ala Ser Phe
20 25 30

```
Phe Cys Val Leu His Asp Arg Gly Asp Ala Pro Ile Asn Thr Pro His
        35                  40                  45
Ala Val Ile Lys Leu Cys Leu Pro Ala Lys Arg Pro Gly Gly Gly Pro
    50                  55                  60
Arg Cys Leu Pro Leu Met Val Leu Asn Leu Pro Ala Trp Gln Val Asn
65                  70                  75                  80
Leu Phe Leu Thr Gly Asp Ala Pro Leu Thr Ser Asp Asn Ile Lys Asp
                85                  90                  95
Arg Ile Asp Leu Ala Gln Thr Glu Glu Ile Leu Glu Pro Ile Leu Ser
            100                 105                 110
Val Leu Ala Cys Lys Arg Ser Ala Gln Gln Thr Lys His Asp Ser Phe
        115                 120                 125
Lys Ser Lys Val Ala Trp Phe Arg Ala Lys Phe Val Ser Ala Leu Arg
    130                 135                 140
Lys Val Tyr Lys Met Thr Pro Ser Pro Tyr Trp Met Ile Thr Leu Leu
145                 150                 155                 160
Gly Ser Phe Glu Ala Ser Phe Val Leu Ala Gly Thr Phe Tyr Phe Phe
                165                 170                 175
Gln Ser His Ile Cys Thr Ala Glu Thr Leu Val His Leu Thr Arg Leu
            180                 185                 190
Phe Ser Ser Ser Gln Gly Gln Ser Leu Val Thr Val Asn Thr Tyr Asp
        195                 200                 205
Glu Leu Gly Arg Val Phe Gly Arg Ser Asp Phe Leu Gly Ile Val Pro
    210                 215                 220
Asn Phe Trp Ala Tyr Leu Lys Tyr Lys Met Gln Gln Asp Asp Val Glu
225                 230                 235                 240
Ser Arg Ala Ile Asp Gln Thr Ile Asn Ser Ile Arg Gly Gly Leu Met
                245                 250                 255
Leu Ser Pro Gln Asp Leu Val His Phe Ile Tyr Leu Ser Phe Tyr Glu
            260                 265                 270
Cys Met Asn Ala Gln Thr Phe Leu Ser Tyr Ser Arg Thr Thr Ser Ser
        275                 280                 285
Leu Pro Thr Pro Ala Thr Val Asn Pro Pro Gln Leu Cys Arg Arg Leu
    290                 295                 300
Glu Ala Asp Phe Lys Glu His Val Met Ala Tyr Tyr Asn Lys Ala Ser
305                 310                 315                 320
Tyr Leu Ser Thr Tyr Ile Thr Ile Leu Thr Val Pro Ala Pro Leu Pro
                325                 330                 335
Asp Gly Tyr Glu Asn Phe Gln Glu Leu Ala Cys Gln Tyr Trp Cys Gly
            340                 345                 350
Gln Ser Arg Asp Val Ala Glu Ile Met Thr Arg Ile Asn Asp Gln Tyr
        355                 360                 365
Pro Gln Leu Asn Leu Thr Lys Asp Leu Ser Gly Leu Leu Asp Leu Ala
    370                 375                 380
Ala Leu Asp Gln Tyr Ser Gly Gly Pro Lys Glu Asn Leu Phe Thr Val
385                 390                 395                 400
Ala Ser Arg Ile Pro Thr Tyr Arg Cys Glu Phe Leu Asn Lys Gln Tyr
                405                 410                 415
Phe Val Leu Met His Ala Asp Cys Ile Asp Ala Tyr Trp Lys Gln Asn
            420                 425                 430
Ile Ile Val Pro Glu Asp Ala Gln Leu Gln Gly Leu Thr Asp Gln Asp
        435                 440                 445
Leu Thr Ser Arg Ile Phe Tyr Cys Asp Leu Gly Leu Ser Leu Pro Thr
    450                 455                 460
Phe Lys Gln Gln Ile Leu Val Ser Arg His Glu Tyr Phe Asn Pro Arg
465                 470                 475                 480
Leu Pro Val Tyr Arg Trp Val Leu Asp Phe Asp Leu Lys Val Thr Glu
                485                 490                 495
Gly Arg Arg Thr Leu Asn Asp Ile Tyr Asn Ile Cys Val Thr Leu Arg
            500                 505                 510
Gln Val Ile Leu Glu Thr Leu Gln Leu Ile Gly Pro Leu Lys Pro Asn
```

515 520 525
His Pro Val Tyr Phe Phe Lys Ser Ala Cys Pro Ala Val Thr Trp Pro
530 535 540
Asp Asp Ile Ser Asp Thr Ala Phe Cys His Cys Asp Ala Lys Ile Gly
545 550 555 560
Met Arg Ile Val Thr Pro Phe Pro Ser Gly Tyr Cys Leu Val Gly Ser
565 570 575
Ala Pro Leu Val Ser Leu Thr Asp Ile Leu Asn Arg Val Val Lys Leu
580 585 590
Asp Thr Arg Leu Ala Ser Glu Tyr Pro Gly Ile Leu Glu Asp Lys Gly
595 600 605
Pro Phe Asp Ser Gly Ile Tyr Ala Lys Gly Arg Cys Val Arg Val Pro
610 615 620
His Cys Tyr Lys Val Gly Pro Gly Gly Glu Leu Ser Arg Leu Leu Lys
625 630 635 640
Ile Ile Ile Cys His Pro Glu Glu Ser Asp Lys Ser Ala Tyr Leu Lys
645 650 655
Asn Ala Phe Lys Val Ser Asn Leu Leu His His Ala Pro Gly Asp Ser
660 665 670
Val Thr Lys Asn Gly His Leu Val Tyr Ala Ile Thr Asp Glu Asn Glu
675 680 685
Gly Phe Leu Glu Ser Lys Thr Lys Asn Asn Leu Pro Lys Thr Ile Thr
690 695 700
Asp Leu Ala Glu Lys Ile Glu Arg Thr Thr Glu Lys Pro Leu Ile Asp
705 710 715 720
Trp Ala Ala Thr Ala Val Trp Pro Lys Leu His Asp Thr Ile Gln Arg
725 730 735
Phe Phe Pro Asp Asp Arg Ile Gly Gln Phe Ala Ser Val Ser Phe Met
740 745 750
His Ser Gly Asp Asn Ile Ile Gln Val Lys Pro Gln Lys Gly Asn Asn
755 760 765
Phe Phe Cys Ile Asn His Lys His Arg Asn His Thr Gln Thr Val Arg
770 775 780
Val Phe Leu Thr Leu His Ser Thr Lys Glu Ser Glu Val Thr Val Thr
785 790 795 800
Phe Met Ser Gln Cys Phe Ala Ala Lys Cys Asn His Asn Ser Pro Thr
805 810 815
Ala His Phe Ser Phe Met Val Pro Ile Thr Gly Thr
820 825

<210> 112
<211> 1329
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1329)

<400> 112
atg cgg tat gtg ttt cac gca cta att tgt ttt ata ggg ggc ata tca 48
Met Arg Tyr Val Phe His Ala Leu Ile Cys Phe Ile Gly Gly Ile Ser
1 5 10 15

tcg tcg gac ttc gac gac tcg tcc tcg gac gaa atg gac gac ctg tcc 96
Ser Ser Asp Phe Asp Asp Ser Ser Ser Asp Glu Met Asp Asp Leu Ser
20 25 30

cca acg ccg gag cca gaa ccg tcc aca acg ccg aac agc ttt ccg gag 144

Pro Thr Pro Glu Pro Glu Pro Ser Thr Thr Pro Asn Ser Phe Pro Glu
35 40 45

ggg ccc aaa tca caa gtg gtg gcg tta cct aaa atc cgt aaa aga tct 192
Gly Pro Lys Ser Gln Val Val Ala Leu Pro Lys Ile Arg Lys Arg Ser
50 55 60

cga tct gag acg ccg gta aaa att gag cac aga tct cca ctt aac cgc 240
Arg Ser Glu Thr Pro Val Lys Ile Glu His Arg Ser Pro Leu Asn Arg
65 70 75 80

tcg cga tct cga tcc aga acg cgg tcc ggg tcc ggt caa cga tca aac 288
Ser Arg Ser Arg Ser Arg Thr Arg Ser Gly Ser Gly Gln Arg Ser Asn
85 90 95

cag tca ggc aga tac gtc aag aga ttc aaa cca acg gtt gat gca ccc 336
Gln Ser Gly Arg Tyr Val Lys Arg Phe Lys Pro Thr Val Asp Ala Pro
100 105 110

cgt cat cga gaa ccg tgg cac agg ggc ggt aag gga aag gcc ccg ttt 384
Arg His Arg Glu Pro Trp His Arg Gly Gly Lys Gly Lys Ala Pro Phe
115 120 125

atc cgc aga gac gca atg gct ggc cgc ggt cga cgc aca tac ggc cac 432
Ile Arg Arg Asp Ala Met Ala Gly Arg Gly Arg Arg Thr Tyr Gly His
130 135 140

gac tat cgc gga aaa gcc gct tta acg cgg agc att aaa gag tct att 480
Asp Tyr Arg Gly Lys Ala Ala Leu Thr Arg Ser Ile Lys Glu Ser Ile
145 150 155 160

aaa aag atg cac ctt cca tcc acc atg ctc tct cgt gcg cac gat aaa 528
Lys Lys Met His Leu Pro Ser Thr Met Leu Ser Arg Ala His Asp Lys
165 170 175

aag gta ttc gag gga ctg ttg cca cga cac ctg gga cag tgc ttt cag 576
Lys Val Phe Glu Gly Leu Leu Pro Arg His Leu Gly Gln Cys Phe Gln
180 185 190

gtg tgc ctg ccg gcg ccg cca ccg ctg caa ccc gag gtg ttc acg gat 624
Val Cys Leu Pro Ala Pro Pro Pro Leu Gln Pro Glu Val Phe Thr Asp
195 200 205

cga cag ctt acc gct ata gtc aag tct ggc ggg cgc aga gac gcg ctg 672
Arg Gln Leu Thr Ala Ile Val Lys Ser Gly Gly Arg Arg Asp Ala Leu
210 215 220

gtg gcc aaa aaa gtt agc ctg gct aaa cta aca agc cta tac aaa ccc 720
Val Ala Lys Lys Val Ser Leu Ala Lys Leu Thr Ser Leu Tyr Lys Pro
225 230 235 240

ctg ctc acg ttt gtg acg ggg aga aac aac cag gcc cac tgg ttg gcg 768
Leu Leu Thr Phe Val Thr Gly Arg Asn Asn Gln Ala His Trp Leu Ala
245 250 255

acg cgc aaa aac acg cta gcg tct gcg gga ctc gag gct ctc gcg gcc 816
Thr Arg Lys Asn Thr Leu Ala Ser Ala Gly Leu Glu Ala Leu Ala Ala
260 265 270

ttt atc gag gaa ggt ctg gcg tgg gcc cag gtg tgc gta tct caa aac 864
Phe Ile Glu Glu Gly Leu Ala Trp Ala Gln Val Cys Val Ser Gln Asn

275 280 285

aga tcg ttg aac gac agc aac cta gat atc att cta gac agt agt cag 912
Arg Ser Leu Asn Asp Ser Asn Leu Asp Ile Ile Leu Asp Ser Ser Gln
290 295 300

agc gtg tgc acc tgg ttc atc tcc aaa atc agg cat ttg cac ata cag 960
Ser Val Cys Thr Trp Phe Ile Ser Lys Ile Arg His Leu His Ile Gln
305 310 315 320

tgc ttc ttg gag aat cag ggt gag gtt agc ctg gtg aag cag cta act 1008
Cys Phe Leu Glu Asn Gln Gly Glu Val Ser Leu Val Lys Gln Leu Thr
325 330 335

tac ctc gtg tgt ata aat aac cgc ctg gcg gag gcg gcg aac ctg gcg 1056
Tyr Leu Val Cys Ile Asn Asn Arg Leu Ala Glu Ala Ala Asn Leu Ala
340 345 350

gga gag gtg aaa ctg aat ttt aaa ctc gga atg ctg att ggg ttc gcc 1104
Gly Glu Val Lys Leu Asn Phe Lys Leu Gly Met Leu Ile Gly Phe Ala
355 360 365

cta acc ttg ccg gcg ctc ctt gca gag cat aag ttg tct ggc gag agc 1152
Leu Thr Leu Pro Ala Leu Leu Ala Glu His Lys Leu Ser Gly Glu Ser
370 375 380

ctg tac ctg ttt agg tct ttc ctg gag aaa tat aga ccg gga gat gtc 1200
Leu Tyr Leu Phe Arg Ser Phe Leu Glu Lys Tyr Arg Pro Gly Asp Val
385 390 395 400

atg gga ctc tta aat tca ata gtg gtg gag cac tac acc aaa tgt cgt 1248
Met Gly Leu Leu Asn Ser Ile Val Val Glu His Tyr Thr Lys Cys Arg
405 410 415

agc gcg gag tgt gtc att aca acc cac gcc atg gta gga tct ggc gaa 1296
Ser Ala Glu Cys Val Ile Thr Thr His Ala Met Val Gly Ser Gly Glu
420 425 430

aac aac aag gga ctt ttt ttc ttt cca gtg taa 1329
Asn Asn Lys Gly Leu Phe Phe Phe Pro Val
435 440

<210> 113
<211> 442
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 113
Met Arg Tyr Val Phe His Ala Leu Ile Cys Phe Ile Gly Gly Ile Ser
1 5 10 15
Ser Ser Asp Phe Asp Asp Ser Ser Ser Asp Glu Met Asp Asp Leu Ser
20 25 30
Pro Thr Pro Glu Pro Glu Pro Ser Thr Thr Pro Asn Ser Phe Pro Glu
35 40 45
Gly Pro Lys Ser Gln Val Val Ala Leu Pro Lys Ile Arg Lys Arg Ser
50 55 60
Arg Ser Glu Thr Pro Val Lys Ile Glu His Arg Ser Pro Leu Asn Arg
65 70 75 80
Ser Arg Ser Arg Ser Arg Thr Arg Ser Gly Ser Gly Gln Arg Ser Asn
85 90 95

Gln Ser Gly Arg Tyr Val Lys Arg Phe Lys Pro Thr Val Asp Ala Pro
100 105 110
Arg His Arg Glu Pro Trp His Arg Gly Gly Lys Gly Lys Ala Pro Phe
115 120 125
Ile Arg Arg Asp Ala Met Ala Gly Arg Gly Arg Arg Thr Tyr Gly His
130 135 140
Asp Tyr Arg Gly Lys Ala Ala Leu Thr Arg Ser Ile Lys Glu Ser Ile
145 150 155 160
Lys Lys Met His Leu Pro Ser Thr Met Leu Ser Arg Ala His Asp Lys
165 170 175
Lys Val Phe Glu Gly Leu Leu Pro Arg His Leu Gly Gln Cys Phe Gln
180 185 190
Val Cys Leu Pro Ala Pro Pro Pro Leu Gln Pro Glu Val Phe Thr Asp
195 200 205
Arg Gln Leu Thr Ala Ile Val Lys Ser Gly Gly Arg Arg Asp Ala Leu
210 215 220
Val Ala Lys Lys Val Ser Leu Ala Lys Leu Thr Ser Leu Tyr Lys Pro
225 230 235 240
Leu Leu Thr Phe Val Thr Gly Arg Asn Asn Gln Ala His Trp Leu Ala
245 250 255
Thr Arg Lys Asn Thr Leu Ala Ser Ala Gly Leu Glu Ala Leu Ala Ala
260 265 270
Phe Ile Glu Glu Gly Leu Ala Trp Ala Gln Val Cys Val Ser Gln Asn
275 280 285
Arg Ser Leu Asn Asp Ser Asn Leu Asp Ile Ile Leu Asp Ser Ser Gln
290 295 300
Ser Val Cys Thr Trp Phe Ile Ser Lys Ile Arg His Leu His Ile Gln
305 310 315 320
Cys Phe Leu Glu Asn Gln Gly Glu Val Ser Leu Val Lys Gln Leu Thr
325 330 335
Tyr Leu Val Cys Ile Asn Asn Arg Leu Ala Glu Ala Ala Asn Leu Ala
340 345 350
Gly Glu Val Lys Leu Asn Phe Lys Leu Gly Met Leu Ile Gly Phe Ala
355 360 365
Leu Thr Leu Pro Ala Leu Leu Ala Glu His Lys Leu Ser Gly Glu Ser
370 375 380
Leu Tyr Leu Phe Arg Ser Phe Leu Glu Lys Tyr Arg Pro Gly Asp Val
385 390 395 400
Met Gly Leu Leu Asn Ser Ile Val Val Glu His Tyr Thr Lys Cys Arg
405 410 415
Ser Ala Glu Cys Val Ile Thr Thr His Ala Met Val Gly Ser Gly Glu
420 425 430
Asn Asn Lys Gly Leu Phe Phe Phe Pro Val
435 440

<210> 114
<211> 1248
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1248)

<400> 114
atg gct acc tgg cgc cca cct cag agc ggt ggt ccg tcc gcc atg gga 48
Met Ala Thr Trp Arg Pro Pro Gln Ser Gly Gly Pro Ser Ala Met Gly
1 5 10 15

```
ctt cga gaa tgg atc gtt act cac gca aat ctt gca acg tat tcg gga    96
Leu Arg Glu Trp Ile Val Thr His Ala Asn Leu Ala Thr Tyr Ser Gly
            20                  25                  30

ctg ttt tgg gca gac gac gaa aaa acc agg gtc gtg ctg gcg act acg   144
Leu Phe Trp Ala Asp Asp Glu Lys Thr Arg Val Val Leu Ala Thr Thr
        35                  40                  45

acc ccg tgg tct gtg gga ttt gat tac ctt cgg gat ggg aag atg tac   192
Thr Pro Trp Ser Val Gly Phe Asp Tyr Leu Arg Asp Gly Lys Met Tyr
    50                  55                  60

gag gac tac tgc aac cag aga aac atc ccc ctt ccg tcc ggg agg tct   240
Glu Asp Tyr Cys Asn Gln Arg Asn Ile Pro Leu Pro Ser Gly Arg Ser
65                  70                  75                  80

agg ctg ggc cag gcc aag gcg cgc cta ttg ggt gcg ata aga aag agt   288
Arg Leu Gly Gln Ala Lys Ala Arg Leu Leu Gly Ala Ile Arg Lys Ser
                85                  90                  95

gcc tat ttc ata gaa gaa aag gac gtt ttg aga cgg agt ttt tca ttt   336
Ala Tyr Phe Ile Glu Glu Lys Asp Val Leu Arg Arg Ser Phe Ser Phe
            100                 105                 110

gca aac gtg gtc ttc agg ttg cgc tcc gat gaa gag atg tta tgc agg   384
Ala Asn Val Val Phe Arg Leu Arg Ser Asp Glu Glu Met Leu Cys Arg
        115                 120                 125

ttg tgc cct cgg gcc tct ggg gta gcg gca gaa ctg cgg ggc ttg agg   432
Leu Cys Pro Arg Ala Ser Gly Val Ala Ala Glu Leu Arg Gly Leu Arg
    130                 135                 140

ttt cgg atg ttt aag cgc aag ggc gcc gac gag gcg ggc agg gta tct   480
Phe Arg Met Phe Lys Arg Lys Gly Ala Asp Glu Ala Gly Arg Val Ser
145                 150                 155                 160

gaa tac acg gta aag cag ctg ctc ggg cta ctt cgt acc cgt ccc gcc   528
Glu Tyr Thr Val Lys Gln Leu Leu Gly Leu Leu Arg Thr Arg Pro Ala
                165                 170                 175

ggt acc ttc acc atg acc gcc ccg gcc acc gag gcg tct gca acc gca   576
Gly Thr Phe Thr Met Thr Ala Pro Ala Thr Glu Ala Ser Ala Thr Ala
            180                 185                 190

acc gcc tcc ggt gaa gac ggt agg cag gac aat agt caa ggg ggc gcg   624
Thr Ala Ser Gly Glu Asp Gly Arg Gln Asp Asn Ser Gln Gly Gly Ala
        195                 200                 205

gtt gcg ttg ccc ggg gaa cac gcc ctg ccg cta tct gca tcg agc ggt   672
Val Ala Leu Pro Gly Glu His Ala Leu Pro Leu Ser Ala Ser Ser Gly
    210                 215                 220

ctt tct gcg tgt ttg gcg cca tct gtt gac gac ccg tgg gga ttt atg   720
Leu Ser Ala Cys Leu Ala Pro Ser Val Asp Asp Pro Trp Gly Phe Met
225                 230                 235                 240

cat att cag gtt tat tat tat ggg gtc ctt caa gca cag act ttt acc   768
His Ile Gln Val Tyr Tyr Tyr Gly Val Leu Gln Ala Gln Thr Phe Thr
                245                 250                 255

cat tct ggg atg ggc gtt cgt tta tcc acg aga cca acg gat aaa aat   816
```

His Ser Gly Met Gly Val Arg Leu Ser Thr Arg Pro Thr Asp Lys Asn
260 265 270

gaa cac cac gtc tgc atg gct ccc ggg ccg ctt cag ttg tgg ttg ccc 864
Glu His His Val Cys Met Ala Pro Gly Pro Leu Gln Leu Trp Leu Pro
275 280 285

ccg gcg cct tat atg gat gac gat ttt atg ttg agc aga ctt gtg aac 912
Pro Ala Pro Tyr Met Asp Asp Asp Phe Met Leu Ser Arg Leu Val Asn
290 295 300

gcg tta cac gcc ctg gag gac ggc att gtt ttg tgc agt tgt cag tat 960
Ala Leu His Ala Leu Glu Asp Gly Ile Val Leu Cys Ser Cys Gln Tyr
305 310 315 320

ggg att atg atg aat ggc tat gga ttt ctt aat tta tgg ttc cgt ggc 1008
Gly Ile Met Met Asn Gly Tyr Gly Phe Leu Asn Leu Trp Phe Arg Gly
325 330 335

aat acc tca aac acc tct gaa cca aga cgc gtg cca tct ggg gtg gga 1056
Asn Thr Ser Asn Thr Ser Glu Pro Arg Arg Val Pro Ser Gly Val Gly
340 345 350

cat cgc gtt ttt gat acc gat gaa tac atg tta aag tta gcc cag tca 1104
His Arg Val Phe Asp Thr Asp Glu Tyr Met Leu Lys Leu Ala Gln Ser
355 360 365

cca cga ccc tct gat ccc ggg cca ccc gat ccg ttt gca cag att tgg 1152
Pro Arg Pro Ser Asp Pro Gly Pro Pro Asp Pro Phe Ala Gln Ile Trp
370 375 380

gtg tct gcc tgg tcc cta tac gag gag gag gat caa tca cag gcg ccg 1200
Val Ser Ala Trp Ser Leu Tyr Glu Glu Glu Asp Gln Ser Gln Ala Pro
385 390 395 400

ata tgt att gtt gtc cac cag cgt gaa ata tat cgg cac ttt gaa tga 1248
Ile Cys Ile Val Val His Gln Arg Glu Ile Tyr Arg His Phe Glu
405 410 415

<210> 115
<211> 415
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 115
Met Ala Thr Trp Arg Pro Pro Gln Ser Gly Gly Pro Ser Ala Met Gly
1 5 10 15
Leu Arg Glu Trp Ile Val Thr His Ala Asn Leu Ala Thr Tyr Ser Gly
20 25 30
Leu Phe Trp Ala Asp Asp Glu Lys Thr Arg Val Val Leu Ala Thr Thr
35 40 45
Thr Pro Trp Ser Val Gly Phe Asp Tyr Leu Arg Asp Gly Lys Met Tyr
50 55 60
Glu Asp Tyr Cys Asn Gln Arg Asn Ile Pro Leu Pro Ser Gly Arg Ser
65 70 75 80
Arg Leu Gly Gln Ala Lys Ala Arg Leu Leu Gly Ala Ile Arg Lys Ser
85 90 95
Ala Tyr Phe Ile Glu Glu Lys Asp Val Leu Arg Arg Ser Phe Ser Phe
100 105 110
Ala Asn Val Val Phe Arg Leu Arg Ser Asp Glu Glu Met Leu Cys Arg

115 120 125
Leu Cys Pro Arg Ala Ser Gly Val Ala Ala Glu Leu Arg Gly Leu Arg
130 135 140
Phe Arg Met Phe Lys Arg Lys Gly Ala Asp Glu Ala Gly Arg Val Ser
145 150 155 160
Glu Tyr Thr Val Lys Gln Leu Leu Gly Leu Leu Arg Thr Arg Pro Ala
165 170 175
Gly Thr Phe Thr Met Thr Ala Pro Ala Thr Glu Ala Ser Ala Thr Ala
180 185 190
Thr Ala Ser Gly Glu Asp Gly Arg Gln Asp Asn Ser Gln Gly Gly Ala
195 200 205
Val Ala Leu Pro Gly Glu His Ala Leu Pro Leu Ser Ala Ser Ser Gly
210 215 220
Leu Ser Ala Cys Leu Ala Pro Ser Val Asp Asp Pro Trp Gly Phe Met
225 230 235 240
His Ile Gln Val Tyr Tyr Tyr Gly Val Leu Gln Ala Gln Thr Phe Thr
245 250 255
His Ser Gly Met Gly Val Arg Leu Ser Thr Arg Pro Thr Asp Lys Asn
260 265 270
Glu His His Val Cys Met Ala Pro Gly Pro Leu Gln Leu Trp Leu Pro
275 280 285
Pro Ala Pro Tyr Met Asp Asp Asp Phe Met Leu Ser Arg Leu Val Asn
290 295 300
Ala Leu His Ala Leu Glu Asp Gly Ile Val Leu Cys Ser Cys Gln Tyr
305 310 315 320
Gly Ile Met Met Asn Gly Tyr Gly Phe Leu Asn Leu Trp Phe Arg Gly
325 330 335
Asn Thr Ser Asn Thr Ser Glu Pro Arg Arg Val Pro Ser Gly Val Gly
340 345 350
His Arg Val Phe Asp Thr Asp Glu Tyr Met Leu Lys Leu Ala Gln Ser
355 360 365
Pro Arg Pro Ser Asp Pro Gly Pro Pro Asp Pro Phe Ala Gln Ile Trp
370 375 380
Val Ser Ala Trp Ser Leu Tyr Glu Glu Glu Asp Gln Ser Gln Ala Pro
385 390 395 400
Ile Cys Ile Val Val His Gln Arg Glu Ile Tyr Arg His Phe Glu
405 410 415

<210> 116
<211> 1248
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1248)

<400> 116
atg gcg ggc cgt gga gtc gat atc agg gcc tgg ctc gtt gcc gcg gtc 48
Met Ala Gly Arg Gly Val Asp Ile Arg Ala Trp Leu Val Ala Ala Val
1 5 10 15

gaa tcg ggc gag tac cgc gga ttg gtt tgg gag aat gag gac aag acc 96
Glu Ser Gly Glu Tyr Arg Gly Leu Val Trp Glu Asn Glu Asp Lys Thr
20 25 30

gtt gtc cgg gtc ccg tgg aac aaa gta act gca gat cgg tcc gta tgg 144
Val Val Arg Val Pro Trp Asn Lys Val Thr Ala Asp Arg Ser Val Trp
35 40 45

aat agc gag aag ttt ttt gat gac tac tgc aac atg cga ggc ata tgt 192
Asn Ser Glu Lys Phe Phe Asp Asp Tyr Cys Asn Met Arg Gly Ile Cys
50 55 60

cag ggc gaa aaa ccg tcc cac tat ggc aga ttc agg aaa atg aga ttc 240
Gln Gly Glu Lys Pro Ser His Tyr Gly Arg Phe Arg Lys Met Arg Phe
65 70 75 80

ttg tac gac atg aga cac cac aaa agt ata cga gag ctt aaa ttt att 288
Leu Tyr Asp Met Arg His His Lys Ser Ile Arg Glu Leu Lys Phe Ile
85 90 95

aat aag gcc tat ggg cga ccg ggt gct cga tat aga ctg ttt cgc ctg 336
Asn Lys Ala Tyr Gly Arg Pro Gly Ala Arg Tyr Arg Leu Phe Arg Leu
100 105 110

ttg ccg gaa ccg gtg gta agt tgt gca atg tgt aat ctg atg tct tcc 384
Leu Pro Glu Pro Val Val Ser Cys Ala Met Cys Asn Leu Met Ser Ser
115 120 125

acc gaa act cag tgc ctt ggt tta atc agt gag ttc caa tac gac cag 432
Thr Glu Thr Gln Cys Leu Gly Leu Ile Ser Glu Phe Gln Tyr Asp Gln
130 135 140

ggg ggt ggg tct gga cgt gaa aga aga cgg gtt ttt acc gca acg gta 480
Gly Gly Gly Ser Gly Arg Glu Arg Arg Arg Val Phe Thr Ala Thr Val
145 150 155 160

ttg gct cgg agt cgg atg gac aaa aac aag cgt gtg cgc gag cat cgg 528
Leu Ala Arg Ser Arg Met Asp Lys Asn Lys Arg Val Arg Glu His Arg
165 170 175

ctg cca ggg gcc atc caa tta acg ttt cta tat ttt gga gag acc gtt 576
Leu Pro Gly Ala Ile Gln Leu Thr Phe Leu Tyr Phe Gly Glu Thr Val
180 185 190

ggc ctc gaa cgt gtt cat gct gga ata cgg gtc tgc agt cga cca tgt 624
Gly Leu Glu Arg Val His Ala Gly Ile Arg Val Cys Ser Arg Pro Cys
195 200 205

ccc gtc ctc gct ggc cac gcg tgt tgt ttc caa gat gag agg act ttg 672
Pro Val Leu Ala Gly His Ala Cys Cys Phe Gln Asp Glu Arg Thr Leu
210 215 220

ttt tta ccg tca ccg ggg gtc gtt gac tgt cag ttt gcc aga gaa gat 720
Phe Leu Pro Ser Pro Gly Val Val Asp Cys Gln Phe Ala Arg Glu Asp
225 230 235 240

cta agg gtt atg cac aaa aag tgc gag aag ggt gtg ttg att acc ctt 768
Leu Arg Val Met His Lys Lys Cys Glu Lys Gly Val Leu Ile Thr Leu
245 250 255

acg gat act ggc att tgc gtg aag aac ctg gag aac cgg gaa atg aag 816
Thr Asp Thr Gly Ile Cys Val Lys Asn Leu Glu Asn Arg Glu Met Lys
260 265 270

gtg ctt act aac aat gag gag gag tat aag gac ctg cca tcc agg caa 864
Val Leu Thr Asn Asn Glu Glu Glu Tyr Lys Asp Leu Pro Ser Arg Gln
275 280 285

ccc gtt cag gtt ttt gat atg gtg gat tat ctt aga gct ctt gca agg 912
Pro Val Gln Val Phe Asp Met Val Asp Tyr Leu Arg Ala Leu Ala Arg
290 295 300

tcc ccg aac cct gga gac gaa cca cct aga gat tat gca cag atc gcc 960
Ser Pro Asn Pro Gly Asp Glu Pro Pro Arg Asp Tyr Ala Gln Ile Ala
305 310 315 320

ctg tgt ctg tcg gtc cag tcg ccg aat ccc gtc gat gcg ccc atc gcc 1008
Leu Cys Leu Ser Val Gln Ser Pro Asn Pro Val Asp Ala Pro Ile Ala
325 330 335

atg aga ctg cgt tac gtg tgc gaa acc tcg tct gta tgc ggt aca gag 1056
Met Arg Leu Arg Tyr Val Cys Glu Thr Ser Ser Val Cys Gly Thr Glu
340 345 350

ggt tgt ttt tac ccc ggg acg act gta aca tca gag ggt cgc acc gac 1104
Gly Cys Phe Tyr Pro Gly Thr Thr Val Thr Ser Glu Gly Arg Thr Asp
355 360 365

tgt tca ttc cag atg gaa gat cca gga gag ggg acg agc caa tcg cat 1152
Cys Ser Phe Gln Met Glu Asp Pro Gly Glu Gly Thr Ser Gln Ser His
370 375 380

gac ccg gca gta gag ctg ggc gac tcc gga cca gat tca atg gac gat 1200
Asp Pro Ala Val Glu Leu Gly Asp Ser Gly Pro Asp Ser Met Asp Asp
385 390 395 400

ccg gac gcg gga acg agc ggt gaa gat gat ggg gtg gcg tgc tcg tga 1248
Pro Asp Ala Gly Thr Ser Gly Glu Asp Asp Gly Val Ala Cys Ser
405 410 415

<210> 117
<211> 415
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 117
Met Ala Gly Arg Gly Val Asp Ile Arg Ala Trp Leu Val Ala Ala Val
1 5 10 15
Glu Ser Gly Glu Tyr Arg Gly Leu Val Trp Glu Asn Glu Asp Lys Thr
20 25 30
Val Val Arg Val Pro Trp Asn Lys Val Thr Ala Asp Arg Ser Val Trp
35 40 45
Asn Ser Glu Lys Phe Phe Asp Asp Tyr Cys Asn Met Arg Gly Ile Cys
50 55 60
Gln Gly Glu Lys Pro Ser His Tyr Gly Arg Phe Arg Lys Met Arg Phe
65 70 75 80
Leu Tyr Asp Met Arg His His Lys Ser Ile Arg Glu Leu Lys Phe Ile
85 90 95
Asn Lys Ala Tyr Gly Arg Pro Gly Ala Arg Tyr Arg Leu Phe Arg Leu
100 105 110
Leu Pro Glu Pro Val Val Ser Cys Ala Met Cys Asn Leu Met Ser Ser
115 120 125
Thr Glu Thr Gln Cys Leu Gly Leu Ile Ser Glu Phe Gln Tyr Asp Gln
130 135 140
Gly Gly Gly Ser Gly Arg Glu Arg Arg Arg Val Phe Thr Ala Thr Val
145 150 155 160
Leu Ala Arg Ser Arg Met Asp Lys Asn Lys Arg Val Arg Glu His Arg
165 170 175

```
Leu Pro Gly Ala Ile Gln Leu Thr Phe Leu Tyr Phe Gly Glu Thr Val
        180                 185                 190
Gly Leu Glu Arg Val His Ala Gly Ile Arg Val Cys Ser Arg Pro Cys
    195                 200                 205
Pro Val Leu Ala Gly His Ala Cys Cys Phe Gln Asp Glu Arg Thr Leu
    210                 215                 220
Phe Leu Pro Ser Pro Gly Val Val Asp Cys Gln Phe Ala Arg Glu Asp
225                 230                 235                 240
Leu Arg Val Met His Lys Lys Cys Glu Lys Gly Val Leu Ile Thr Leu
                245                 250                 255
Thr Asp Thr Gly Ile Cys Val Lys Asn Leu Glu Asn Arg Glu Met Lys
            260                 265                 270
Val Leu Thr Asn Asn Glu Glu Glu Tyr Lys Asp Leu Pro Ser Arg Gln
        275                 280                 285
Pro Val Gln Val Phe Asp Met Val Asp Tyr Leu Arg Ala Leu Ala Arg
    290                 295                 300
Ser Pro Asn Pro Gly Asp Glu Pro Pro Arg Asp Tyr Ala Gln Ile Ala
305                 310                 315                 320
Leu Cys Leu Ser Val Gln Ser Pro Asn Pro Val Asp Ala Pro Ile Ala
                325                 330                 335
Met Arg Leu Arg Tyr Val Cys Glu Thr Ser Ser Val Cys Gly Thr Glu
            340                 345                 350
Gly Cys Phe Tyr Pro Gly Thr Thr Val Thr Ser Glu Gly Arg Thr Asp
        355                 360                 365
Cys Ser Phe Gln Met Glu Asp Pro Gly Glu Gly Thr Ser Gln Ser His
    370                 375                 380
Asp Pro Ala Val Glu Leu Gly Asp Ser Gly Pro Asp Ser Met Asp Asp
385                 390                 395                 400
Pro Asp Ala Gly Thr Ser Gly Glu Asp Asp Gly Val Ala Cys Ser
                405                 410                 415


<210> 118
<211> 1056
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1056)

<400> 118
atg gag cgg ccg gtc cga gtg acc aaa cca tcc agc ctt cga gga tgg      48
Met Glu Arg Pro Val Arg Val Thr Lys Pro Ser Ser Leu Arg Gly Trp
1               5                   10                  15

ctg gtg gaa tgc tgc gag aca ggt cgt cat ccc gga atg cgg tgg gtt      96
Leu Val Glu Cys Cys Glu Thr Gly Arg His Pro Gly Met Arg Trp Val
            20                  25                  30

gac gag gaa aga act tta atc cgc ata ccg tgg aat cac gac aga ggg     144
Asp Glu Glu Arg Thr Leu Ile Arg Ile Pro Trp Asn His Asp Arg Gly
        35                  40                  45

agc aga gga gtc gag gag gcg gaa aaa aac att ttt att gac tac tgt     192
Ser Arg Gly Val Glu Glu Ala Glu Lys Asn Ile Phe Ile Asp Tyr Cys
    50                  55                  60

cga tcc agg ggc atc ttg cac gcg gct ggc agg gaa tta acc gct aag     240
Arg Ser Arg Gly Ile Leu His Ala Ala Gly Arg Glu Leu Thr Ala Lys
```

```
65                  70                  75                  80

gaa tgc aaa aac tgg cta tcc agc gcc ata cgg cat agt caa acc gtg     288
Glu Cys Lys Asn Trp Leu Ser Ser Ala Ile Arg His Ser Gln Thr Val
                85                  90                  95

tcg gac gtg tca act aag gac aac ctt tcg act cca tat ccg gac aga     336
Ser Asp Val Ser Thr Lys Asp Asn Leu Ser Thr Pro Tyr Pro Asp Arg
            100                 105                 110

tgc cga atc att cgc ctt tta cca att acc gtg agg tcg tgt gct cgt     384
Cys Arg Ile Ile Arg Leu Leu Pro Ile Thr Val Arg Ser Cys Ala Arg
        115                 120                 125

tgc gat cag gcc tcc ggc act act gct atg tta cgg ggg ttg cgc gag     432
Cys Asp Gln Ala Ser Gly Thr Thr Ala Met Leu Arg Gly Leu Arg Glu
    130                 135                 140

gag gcc gtg aat aag ttt ggt ccg gtt ggc gcc ggt gtt cgg tac act     480
Glu Ala Val Asn Lys Phe Gly Pro Val Gly Ala Gly Val Arg Tyr Thr
145                 150                 155                 160

ggt gcg gtg gga gcc ggg ggt gaa cag tgt tgg atg cta cgt ata atg     528
Gly Ala Val Gly Ala Gly Gly Glu Gln Cys Trp Met Leu Arg Ile Met
                165                 170                 175

ttt tac tat tat ggg gac aga gtt ggt gaa gtg gtg act gaa tcg tcc     576
Phe Tyr Tyr Tyr Gly Asp Arg Val Gly Glu Val Val Thr Glu Ser Ser
            180                 185                 190

aac ggc atc cgt gtg ctc cca ttg tcc gaa cgt cgc cct caa ggg cac     624
Asn Gly Ile Arg Val Leu Pro Leu Ser Glu Arg Arg Pro Gln Gly His
        195                 200                 205

ata tgc gcc gca ccg atc gcc gag cag gca ctg gtt ccg gag att cca     672
Ile Cys Ala Ala Pro Ile Ala Glu Gln Ala Leu Val Pro Glu Ile Pro
    210                 215                 220

ggc cac tta gca gag ttt cag gcc gag gcg ttg cgc ttc ctg gat aaa     720
Gly His Leu Ala Glu Phe Gln Ala Glu Ala Leu Arg Phe Leu Asp Lys
225                 230                 235                 240

gat ctt ttg cgc ggg ttg gcg ttt tgg gcc gac ccc tcc ggc att tac     768
Asp Leu Leu Arg Gly Leu Ala Phe Trp Ala Asp Pro Ser Gly Ile Tyr
                245                 250                 255

att aga tgg ttg ggg cat tcc ctg gcg ttc gtt caa ggg aac gtc gaa     816
Ile Arg Trp Leu Gly His Ser Leu Ala Phe Val Gln Gly Asn Val Glu
            260                 265                 270

tct acg ggg gcg gtt gcg gtc ctg tcg tgc gcc aac gcg tgt cgt gca     864
Ser Thr Gly Ala Val Ala Val Leu Ser Cys Ala Asn Ala Cys Arg Ala
        275                 280                 285

ttt aac ttg gtt gac tat atg aca gcc atg gcc aga aca tct ccg gac     912
Phe Asn Leu Val Asp Tyr Met Thr Ala Met Ala Arg Thr Ser Pro Asp
    290                 295                 300

gga gcg gcg ctt cca caa gct tgt gtt tac cta tat ttt ggg ggc gta     960
Gly Ala Ala Leu Pro Gln Ala Cys Val Tyr Leu Tyr Phe Gly Gly Val
305                 310                 315                 320
```

cct aca ccc gag ggc gga gtg caa agc act gta ccg cta att att cag 1008
Pro Thr Pro Glu Gly Gly Val Gln Ser Thr Val Pro Leu Ile Ile Gln
325 330 335

ctg tgg cac gag tgt ctg tgg cgg gcg ttg tcc gcc gct aat gtt taa 1056
Leu Trp His Glu Cys Leu Trp Arg Ala Leu Ser Ala Ala Asn Val
340 345 350

<210> 119
<211> 351
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 119
Met Glu Arg Pro Val Arg Val Thr Lys Pro Ser Ser Leu Arg Gly Trp
1 5 10 15
Leu Val Glu Cys Cys Glu Thr Gly Arg His Pro Gly Met Arg Trp Val
20 25 30
Asp Glu Glu Arg Thr Leu Ile Arg Ile Pro Trp Asn His Asp Arg Gly
35 40 45
Ser Arg Gly Val Glu Glu Ala Glu Lys Asn Ile Phe Ile Asp Tyr Cys
50 55 60
Arg Ser Arg Gly Ile Leu His Ala Ala Gly Arg Glu Leu Thr Ala Lys
65 70 75 80
Glu Cys Lys Asn Trp Leu Ser Ser Ala Ile Arg His Ser Gln Thr Val
85 90 95
Ser Asp Val Ser Thr Lys Asp Asn Leu Ser Thr Pro Tyr Pro Asp Arg
100 105 110
Cys Arg Ile Ile Arg Leu Leu Pro Ile Thr Val Arg Ser Cys Ala Arg
115 120 125
Cys Asp Gln Ala Ser Gly Thr Thr Ala Met Leu Arg Gly Leu Arg Glu
130 135 140
Glu Ala Val Asn Lys Phe Gly Pro Val Gly Ala Gly Val Arg Tyr Thr
145 150 155 160
Gly Ala Val Gly Ala Gly Gly Glu Gln Cys Trp Met Leu Arg Ile Met
165 170 175
Phe Tyr Tyr Tyr Gly Asp Arg Val Gly Glu Val Val Thr Glu Ser Ser
180 185 190
Asn Gly Ile Arg Val Leu Pro Leu Ser Glu Arg Arg Pro Gln Gly His
195 200 205
Ile Cys Ala Ala Pro Ile Ala Glu Gln Ala Leu Val Pro Glu Ile Pro
210 215 220
Gly His Leu Ala Glu Phe Gln Ala Glu Ala Leu Arg Phe Leu Asp Lys
225 230 235 240
Asp Leu Leu Arg Gly Leu Ala Phe Trp Ala Asp Pro Ser Gly Ile Tyr
245 250 255
Ile Arg Trp Leu Gly His Ser Leu Ala Phe Val Gln Gly Asn Val Glu
260 265 270
Ser Thr Gly Ala Val Ala Val Leu Ser Cys Ala Asn Ala Cys Arg Ala
275 280 285
Phe Asn Leu Val Asp Tyr Met Thr Ala Met Ala Arg Thr Ser Pro Asp
290 295 300
Gly Ala Ala Leu Pro Gln Ala Cys Val Tyr Leu Tyr Phe Gly Gly Val
305 310 315 320
Pro Thr Pro Glu Gly Gly Val Gln Ser Thr Val Pro Leu Ile Ile Gln
325 330 335
Leu Trp His Glu Cys Leu Trp Arg Ala Leu Ser Ala Ala Asn Val
340 345 350

```
<210> 120
<211> 762
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(762)

<400> 120
atg gat tct gga tgc tat gcg tgt atc ctt gat gag aac tct gaa ggg      48
Met Asp Ser Gly Cys Tyr Ala Cys Ile Leu Asp Glu Asn Ser Glu Gly
1               5                   10                  15

att ata aat tat ttg gag cag gtg tgt ggc atc ggg tta gag ccg gga      96
Ile Ile Asn Tyr Leu Glu Gln Val Cys Gly Ile Gly Leu Glu Pro Gly
            20                  25                  30

atg ccg ctg ccc gcc ccc ttg ccc acg ttg gtc cct ccg acg aga tca      144
Met Pro Leu Pro Ala Pro Leu Pro Thr Leu Val Pro Pro Thr Arg Ser
        35                  40                  45

gcg tac gcc cga gca cat cgt ctc ggg gtc cca gaa gcg ccg ctg cct      192
Ala Tyr Ala Arg Ala His Arg Leu Gly Val Pro Glu Ala Pro Leu Pro
    50                  55                  60

cat caa att gtg ccg ttt tgg cgg ttg aga att cag gtg ttt tat ttt      240
His Gln Ile Val Pro Phe Trp Arg Leu Arg Ile Gln Val Phe Tyr Phe
65                  70                  75                  80

ggc gtt ctg gca ctg gat cac acg tcc caa gat cgg aga ggt gtt cgt      288
Gly Val Leu Ala Leu Asp His Thr Ser Gln Asp Arg Arg Gly Val Arg
                85                  90                  95

ttg cat ccg cgc ccg gtt ccc cat ccc ggt cat ctg tgt ttc tat ggc      336
Leu His Pro Arg Pro Val Pro His Pro Gly His Leu Cys Phe Tyr Gly
            100                 105                 110

acc ggg ttc act gtc tgg ttt cca tct cca gat cgt gaa aaa ctg acc      384
Thr Gly Phe Thr Val Trp Phe Pro Ser Pro Asp Arg Glu Lys Leu Thr
        115                 120                 125

gca gaa cag ata acg caa ata aaa acg atg ctg gtc gcg tat aac gag      432
Ala Glu Gln Ile Thr Gln Ile Lys Thr Met Leu Val Ala Tyr Asn Glu
    130                 135                 140

ggc atc tac gta cac ggc aac gaa acc gga gtg tac gtc gat aat agg      480
Gly Ile Tyr Val His Gly Asn Glu Thr Gly Val Tyr Val Asp Asn Arg
145                 150                 155                 160

aac agg gaa acg ctg tac gcg gcg ggt aac gac tgt aat ggt gat att      528
Asn Arg Glu Thr Leu Tyr Ala Ala Gly Asn Asp Cys Asn Gly Asp Ile
                165                 170                 175

atc caa cga gag gtc atg ttc ctc tct aaa cag aaa att ttt aac ttc      576
Ile Gln Arg Glu Val Met Phe Leu Ser Lys Gln Lys Ile Phe Asn Phe
            180                 185                 190

atg ggt ttc atg aga aag ctt gcc cgc tct ccg gga cca gag agc cat      624
```

Met Gly Phe Met Arg Lys Leu Ala Arg Ser Pro Gly Pro Glu Ser His
195 200 205

gcg ccg tgc aac ggg gca aca ctg tat ctc tca cag cag cct ggt gcc 672
Ala Pro Cys Asn Gly Ala Thr Leu Tyr Leu Ser Gln Gln Pro Gly Ala
210 215 220

caa gag agc cca caa gtt cct att tct gtg gtg gtc tgt cag gac gag 720
Gln Glu Ser Pro Gln Val Pro Ile Ser Val Val Val Cys Gln Asp Glu
225 230 235 240

ctg gtg cag gga cac atg aat cct tcc cgg tgg tgc tca taa 762
Leu Val Gln Gly His Met Asn Pro Ser Arg Trp Cys Ser
245 250

<210> 121
<211> 253
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 121
Met Asp Ser Gly Cys Tyr Ala Cys Ile Leu Asp Glu Asn Ser Glu Gly
1 5 10 15
Ile Ile Asn Tyr Leu Glu Gln Val Cys Gly Ile Gly Leu Glu Pro Gly
20 25 30
Met Pro Leu Pro Ala Pro Leu Pro Thr Leu Val Pro Pro Thr Arg Ser
35 40 45
Ala Tyr Ala Arg Ala His Arg Leu Gly Val Pro Glu Ala Pro Leu Pro
50 55 60
His Gln Ile Val Pro Phe Trp Arg Leu Arg Ile Gln Val Phe Tyr Phe
65 70 75 80
Gly Val Leu Ala Leu Asp His Thr Ser Gln Asp Arg Arg Gly Val Arg
85 90 95
Leu His Pro Arg Pro Val Pro His Pro Gly His Leu Cys Phe Tyr Gly
100 105 110
Thr Gly Phe Thr Val Trp Phe Pro Ser Pro Asp Arg Glu Lys Leu Thr
115 120 125
Ala Glu Gln Ile Thr Gln Ile Lys Thr Met Leu Val Ala Tyr Asn Glu
130 135 140
Gly Ile Tyr Val His Gly Asn Glu Thr Gly Val Tyr Val Asp Asn Arg
145 150 155 160
Asn Arg Glu Thr Leu Tyr Ala Ala Gly Asn Asp Cys Asn Gly Asp Ile
165 170 175
Ile Gln Arg Glu Val Met Phe Leu Ser Lys Gln Lys Ile Phe Asn Phe
180 185 190
Met Gly Phe Met Arg Lys Leu Ala Arg Ser Pro Gly Pro Glu Ser His
195 200 205
Ala Pro Cys Asn Gly Ala Thr Leu Tyr Leu Ser Gln Gln Pro Gly Ala
210 215 220
Gln Glu Ser Pro Gln Val Pro Ile Ser Val Val Val Cys Gln Asp Glu
225 230 235 240
Leu Val Gln Gly His Met Asn Pro Ser Arg Trp Cys Ser
245 250

<210> 122
<211> 1158
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1158)

<400> 122
atg gcc gct ggg gaa tcg aga agg ggc cca tct cgt tat ggc atg gcc 48
Met Ala Ala Gly Glu Ser Arg Arg Gly Pro Ser Arg Tyr Gly Met Ala
1 5 10 15

ctc aag gaa tgg ctg acg ttc aaa gcc gat agt ggt ctg tac ccg ggc 96
Leu Lys Glu Trp Leu Thr Phe Lys Ala Asp Ser Gly Leu Tyr Pro Gly
20 25 30

ctg ttt tgg gca gac gag caa aag acg agg ctg gtg ttg gca gct aca 144
Leu Phe Trp Ala Asp Glu Gln Lys Thr Arg Leu Val Leu Ala Ala Thr
35 40 45

cca ccg tcg ttt ccc aat tat gat tat caa cgc gac ggt caa cac tat 192
Pro Pro Ser Phe Pro Asn Tyr Asp Tyr Gln Arg Asp Gly Gln His Tyr
50 55 60

gat gcc tac tgt gag tta aga cac atc cca ctg ccg tct gga agg agc 240
Asp Ala Tyr Cys Glu Leu Arg His Ile Pro Leu Pro Ser Gly Arg Ser
65 70 75 80

agg ctg tgt caa gcc agg ggg cgc ctg ctg gga gcc gta aga aag agc 288
Arg Leu Cys Gln Ala Arg Gly Arg Leu Leu Gly Ala Val Arg Lys Ser
85 90 95

aag tac ttt gag gaa gat aag gaa ttc cct acc gat caa ttt ccg ttc 336
Lys Tyr Phe Glu Glu Asp Lys Glu Phe Pro Thr Asp Gln Phe Pro Phe
100 105 110

acg gcg ctg gtc ttt agg ttg cgt tca tcc gaa gag atg tct tgt ccg 384
Thr Ala Leu Val Phe Arg Leu Arg Ser Ser Glu Glu Met Ser Cys Pro
115 120 125

gtg tgt ccg cgc gtg tgc gcg ctc cgc ctt gag ttg cgc aat atg cgg 432
Val Cys Pro Arg Val Cys Ala Leu Arg Leu Glu Leu Arg Asn Met Arg
130 135 140

ttt gcc atg ctc ggg cgc gga atg ctt cac gcc ctt tcg gga ccg tcc 480
Phe Ala Met Leu Gly Arg Gly Met Leu His Ala Leu Ser Gly Pro Ser
145 150 155 160

gta agt gat caa gag aga cgg tat cgg gaa ggt cat caa gat gga cac 528
Val Ser Asp Gln Glu Arg Arg Tyr Arg Glu Gly His Gln Asp Gly His
165 170 175

gac gcg cag gac gac gat gca gcg tat tct tcg ggc ctt ctg cgc gcc 576
Asp Ala Gln Asp Asp Asp Ala Ala Tyr Ser Ser Gly Leu Leu Arg Ala
180 185 190

aga ctt atg gcg tgt gct gcg cca tct gcc ggc gat cca tgg ggt cac 624
Arg Leu Met Ala Cys Ala Ala Pro Ser Ala Gly Asp Pro Trp Gly His
195 200 205

atg cat ata aag att tat tac tat ggg caa ttg cag gct gag ctg ttg 672
Met His Ile Lys Ile Tyr Tyr Tyr Gly Gln Leu Gln Ala Glu Leu Leu
210 215 220

aca gcc acc ggg cag gga atc agg cta tct tca aaa cca aca aat aag 720
Thr Ala Thr Gly Gln Gly Ile Arg Leu Ser Ser Lys Pro Thr Asn Lys
225 230 235 240

gcg gga cac cac gtg tgt gtt ttg gat gga ccc ctt cag gcg tgg ttt 768
Ala Gly His His Val Cys Val Leu Asp Gly Pro Leu Gln Ala Trp Phe
245 250 255

cct ccg ata cca cag acg act gag tcg tct gtg gtg cag aga ctc gag 816
Pro Pro Ile Pro Gln Thr Thr Glu Ser Ser Val Val Gln Arg Leu Glu
260 265 270

gac gct ttg aaa tgg ctc gtg gat gga ata att ttc tgc agc acc agc 864
Asp Ala Leu Lys Trp Leu Val Asp Gly Ile Ile Phe Cys Ser Thr Ser
275 280 285

agg gga att atg ttc acg att acc ggg ggt cct aac gtg tgg ttt cag 912
Arg Gly Ile Met Phe Thr Ile Thr Gly Gly Pro Asn Val Trp Phe Gln
290 295 300

gga aat acc gtg gaa ccg tac agt ttg cca cat cgt gcc tac acc ggg 960
Gly Asn Thr Val Glu Pro Tyr Ser Leu Pro His Arg Ala Tyr Thr Gly
305 310 315 320

atg cac gtg tgg gca ttt gac act gat cgt tat ttg ctt gat atg gcc 1008
Met His Val Trp Ala Phe Asp Thr Asp Arg Tyr Leu Leu Asp Met Ala
325 330 335

agg tcc ccc tcc ccg cgc gac acg ggt ccc ccc gcg gcc ttt gtt aaa 1056
Arg Ser Pro Ser Pro Arg Asp Thr Gly Pro Pro Ala Ala Phe Val Lys
340 345 350

ctc tgg gtt tcg ggt tgt tcc ctg ggc gag gag cga aac tct tcg agg 1104
Leu Trp Val Ser Gly Cys Ser Leu Gly Glu Glu Arg Asn Ser Ser Arg
355 360 365

gct cca cta tct att att gtt tat caa aca gaa atc tat agg cac ttt 1152
Ala Pro Leu Ser Ile Ile Val Tyr Gln Thr Glu Ile Tyr Arg His Phe
370 375 380

gaa taa 1158
Glu
385

<210> 123
<211> 385
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 123
Met Ala Ala Gly Glu Ser Arg Arg Gly Pro Ser Arg Tyr Gly Met Ala
1 5 10 15
Leu Lys Glu Trp Leu Thr Phe Lys Ala Asp Ser Gly Leu Tyr Pro Gly
20 25 30
Leu Phe Trp Ala Asp Glu Gln Lys Thr Arg Leu Val Leu Ala Ala Thr
35 40 45
Pro Pro Ser Phe Pro Asn Tyr Asp Tyr Gln Arg Asp Gly Gln His Tyr
50 55 60
Asp Ala Tyr Cys Glu Leu Arg His Ile Pro Leu Pro Ser Gly Arg Ser

65 70 75 80
Arg Leu Cys Gln Ala Arg Gly Arg Leu Leu Gly Ala Val Arg Lys Ser
85 90 95
Lys Tyr Phe Glu Glu Asp Lys Glu Phe Pro Thr Asp Gln Phe Pro Phe
100 105 110
Thr Ala Leu Val Phe Arg Leu Arg Ser Ser Glu Glu Met Ser Cys Pro
115 120 125
Val Cys Pro Arg Val Cys Ala Leu Arg Leu Glu Leu Arg Asn Met Arg
130 135 140
Phe Ala Met Leu Gly Arg Gly Met Leu His Ala Leu Ser Gly Pro Ser
145 150 155 160
Val Ser Asp Gln Glu Arg Arg Tyr Arg Glu Gly His Gln Asp Gly His
165 170 175
Asp Ala Gln Asp Asp Asp Ala Ala Tyr Ser Ser Gly Leu Leu Arg Ala
180 185 190
Arg Leu Met Ala Cys Ala Ala Pro Ser Ala Gly Asp Pro Trp Gly His
195 200 205
Met His Ile Lys Ile Tyr Tyr Tyr Gly Gln Leu Gln Ala Glu Leu Leu
210 215 220
Thr Ala Thr Gly Gln Gly Ile Arg Leu Ser Ser Lys Pro Thr Asn Lys
225 230 235 240
Ala Gly His His Val Cys Val Leu Asp Gly Pro Leu Gln Ala Trp Phe
245 250 255
Pro Pro Ile Pro Gln Thr Thr Glu Ser Ser Val Val Gln Arg Leu Glu
260 265 270
Asp Ala Leu Lys Trp Leu Val Asp Gly Ile Ile Phe Cys Ser Thr Ser
275 280 285
Arg Gly Ile Met Phe Thr Ile Thr Gly Gly Pro Asn Val Trp Phe Gln
290 295 300
Gly Asn Thr Val Glu Pro Tyr Ser Leu Pro His Arg Ala Tyr Thr Gly
305 310 315 320
Met His Val Trp Ala Phe Asp Thr Asp Arg Tyr Leu Leu Asp Met Ala
325 330 335
Arg Ser Pro Ser Pro Arg Asp Thr Gly Pro Pro Ala Ala Phe Val Lys
340 345 350
Leu Trp Val Ser Gly Cys Ser Leu Gly Glu Glu Arg Asn Ser Ser Arg
355 360 365
Ala Pro Leu Ser Ile Ile Val Tyr Gln Thr Glu Ile Tyr Arg His Phe
370 375 380
Glu
385

<210> 124
<211> 1173
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1173)

<400> 124
atg gcg gaa cgc gat atg gat tta aaa gct tgg ttc atc gag gcc gtg 48
Met Ala Glu Arg Asp Met Asp Leu Lys Ala Trp Phe Ile Glu Ala Val
1 5 10 15

gag tct aag cga tac ccc gga gtg gaa tgg gat gac gag gac aag acc 96
Glu Ser Lys Arg Tyr Pro Gly Val Glu Trp Asp Asp Glu Asp Lys Thr
20 25 30

att ata cgt gtc ccg tgg aac cgg tgc acc gac agt cgc gtc gat gaa 144
Ile Ile Arg Val Pro Trp Asn Arg Cys Thr Asp Ser Arg Val Asp Glu
35 40 45

gat tat aac aag ata ttt gac gat ttt tgc tca gcg cgt ggc gtg tgt 192
Asp Tyr Asn Lys Ile Phe Asp Asp Phe Cys Ser Ala Arg Gly Val Cys
50 55 60

caa acc ggt tca cat gcc cag aaa ttt aag aaa att agg atg tta tat 240
Gln Thr Gly Ser His Ala Gln Lys Phe Lys Lys Ile Arg Met Leu Tyr
65 70 75 80

gca gtg aga tct cac agg tat ttg agg gag ctg aca ccg ccg agc aag 288
Ala Val Arg Ser His Arg Tyr Leu Arg Glu Leu Thr Pro Pro Ser Lys
85 90 95

gcc ggg ggc gtc tct ggg gaa aga tac aga ctc ttt caa ttg ctt cct 336
Ala Gly Gly Val Ser Gly Glu Arg Tyr Arg Leu Phe Gln Leu Leu Pro
100 105 110

gag gtt acg gtg ggc tgc gat ctg tgt aac ctc atc gcg acc aca tcg 384
Glu Val Thr Val Gly Cys Asp Leu Cys Asn Leu Ile Ala Thr Thr Ser
115 120 125

ttg cat agc tgt tcc atg ggc agt tgc gtt cga gag gat gtt ttc gag 432
Leu His Ser Cys Ser Met Gly Ser Cys Val Arg Glu Asp Val Phe Glu
130 135 140

agg aca cgg cgg ccg agg gct aag gcg cca ctg aga gtc tcc gtt tat 480
Arg Thr Arg Arg Pro Arg Ala Lys Ala Pro Leu Arg Val Ser Val Tyr
145 150 155 160

aaa cgg aaa tcc aaa cga ctc cag gac tct tct gcg caa cct gtg cta 528
Lys Arg Lys Ser Lys Arg Leu Gln Asp Ser Ser Ala Gln Pro Val Leu
165 170 175

ggg gct gtt gag gta tcc ttt ttt tac ttt gga gag aac gtc gga gtt 576
Gly Ala Val Glu Val Ser Phe Phe Tyr Phe Gly Glu Asn Val Gly Val
180 185 190

caa att ttg cga gcc ggc tcg ggt gtg cgt ata tgt ggc ctt cct gat 624
Gln Ile Leu Arg Ala Gly Ser Gly Val Arg Ile Cys Gly Leu Pro Asp
195 200 205

ccc aaa cgt ccc gga cac ctg tgt tgt gcc gat aat cca ttg acg tgt 672
Pro Lys Arg Pro Gly His Leu Cys Cys Ala Asp Asn Pro Leu Thr Cys
210 215 220

ttc ttg ccg tcg tcc cag ttg att ccc tgc gag ttt gcc agg gca gat 720
Phe Leu Pro Ser Ser Gln Leu Ile Pro Cys Glu Phe Ala Arg Ala Asp
225 230 235 240

ttg cag gcc ctg cag aaa aca tgc gag aga ggt ttg atc tgt gtt atg 768
Leu Gln Ala Leu Gln Lys Thr Cys Glu Arg Gly Leu Ile Cys Val Met
245 250 255

acg gag tct gga atc tgt gtg aaa aac ctg gaa gaa cgc aac atg acg 816
Thr Glu Ser Gly Ile Cys Val Lys Asn Leu Glu Glu Arg Asn Met Thr
260 265 270

gca ttg acc aac tat tct gaa aat tac tat gag ctg cgg cct tcg cag 864
Ala Leu Thr Asn Tyr Ser Glu Asn Tyr Tyr Glu Leu Arg Pro Ser Gln
275 280 285

ccg ctc cag gcg ttc gat ctc ttg cat tat ttg agg gag ctg gcg aga 912
Pro Leu Gln Ala Phe Asp Leu Leu His Tyr Leu Arg Glu Leu Ala Arg
290 295 300

tcg cca act cca gga gat gtc ccg ccg cgt gac tgt gcg tgg att ttc 960
Ser Pro Thr Pro Gly Asp Val Pro Pro Arg Asp Cys Ala Trp Ile Phe
305 310 315 320

atg tgc ccg tct act cag tcc gaa aat aca tgg gat gct ccc att gcc 1008
Met Cys Pro Ser Thr Gln Ser Glu Asn Thr Trp Asp Ala Pro Ile Ala
325 330 335

ctg aaa ctt cgc tac gtg tgc aat gat gat gtg tca gat gat gtc agc 1056
Leu Lys Leu Arg Tyr Val Cys Asn Asp Asp Val Ser Asp Asp Val Ser
340 345 350

aac ggt gcc gcg gga gat gac agc ggg gac gag gga ccg tct gga gcg 1104
Asn Gly Ala Ala Gly Asp Asp Ser Gly Asp Glu Gly Pro Ser Gly Ala
355 360 365

ggt gtc ggt gct tcg gga aca acg gga agc aca tct gta tct acc ctc 1152
Gly Val Gly Ala Ser Gly Thr Thr Gly Ser Thr Ser Val Ser Thr Leu
370 375 380

gca ccg tat ggg agg aag taa 1173
Ala Pro Tyr Gly Arg Lys
385 390

<210> 125
<211> 390
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 125
Met Ala Glu Arg Asp Met Asp Leu Lys Ala Trp Phe Ile Glu Ala Val
1 5 10 15
Glu Ser Lys Arg Tyr Pro Gly Val Glu Trp Asp Asp Glu Asp Lys Thr
20 25 30
Ile Ile Arg Val Pro Trp Asn Arg Cys Thr Asp Ser Arg Val Asp Glu
35 40 45
Asp Tyr Asn Lys Ile Phe Asp Asp Phe Cys Ser Ala Arg Gly Val Cys
50 55 60
Gln Thr Gly Ser His Ala Gln Lys Phe Lys Lys Ile Arg Met Leu Tyr
65 70 75 80
Ala Val Arg Ser His Arg Tyr Leu Arg Glu Leu Thr Pro Pro Ser Lys
85 90 95
Ala Gly Gly Val Ser Gly Glu Arg Tyr Arg Leu Phe Gln Leu Leu Pro
100 105 110
Glu Val Thr Val Gly Cys Asp Leu Cys Asn Leu Ile Ala Thr Thr Ser
115 120 125
Leu His Ser Cys Ser Met Gly Ser Cys Val Arg Glu Asp Val Phe Glu
130 135 140
Arg Thr Arg Arg Pro Arg Ala Lys Ala Pro Leu Arg Val Ser Val Tyr
145 150 155 160
Lys Arg Lys Ser Lys Arg Leu Gln Asp Ser Ser Ala Gln Pro Val Leu
165 170 175

Gly Ala Val Glu Val Ser Phe Phe Tyr Phe Gly Glu Asn Val Gly Val
180 185 190
Gln Ile Leu Arg Ala Gly Ser Gly Val Arg Ile Cys Gly Leu Pro Asp
195 200 205
Pro Lys Arg Pro Gly His Leu Cys Cys Ala Asp Asn Pro Leu Thr Cys
210 215 220
Phe Leu Pro Ser Ser Gln Leu Ile Pro Cys Glu Phe Ala Arg Ala Asp
225 230 235 240
Leu Gln Ala Leu Gln Lys Thr Cys Glu Arg Gly Leu Ile Cys Val Met
245 250 255
Thr Glu Ser Gly Ile Cys Val Lys Asn Leu Glu Glu Arg Asn Met Thr
260 265 270
Ala Leu Thr Asn Tyr Ser Glu Asn Tyr Tyr Glu Leu Arg Pro Ser Gln
275 280 285
Pro Leu Gln Ala Phe Asp Leu Leu His Tyr Leu Arg Glu Leu Ala Arg
290 295 300
Ser Pro Thr Pro Gly Asp Val Pro Pro Arg Asp Cys Ala Trp Ile Phe
305 310 315 320
Met Cys Pro Ser Thr Gln Ser Glu Asn Thr Trp Asp Ala Pro Ile Ala
325 330 335
Leu Lys Leu Arg Tyr Val Cys Asn Asp Asp Val Ser Asp Asp Val Ser
340 345 350
Asn Gly Ala Ala Gly Asp Asp Ser Gly Asp Glu Gly Pro Ser Gly Ala
355 360 365
Gly Val Gly Ala Ser Gly Thr Thr Gly Ser Thr Ser Val Ser Thr Leu
370 375 380
Ala Pro Tyr Gly Arg Lys
385 390

<210> 126
<211> 1068
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1068)

<400> 126
atg gcg gaa gga cgg gca ggg tcg atc cgc gta aac cgg ccg tct ggc 48
Met Ala Glu Gly Arg Ala Gly Ser Ile Arg Val Asn Arg Pro Ser Gly
1 5 10 15

ctg agg gca tgg ttg ctc gat tgt tgc gat aat gat aag cat cct gga 96
Leu Arg Ala Trp Leu Leu Asp Cys Cys Asp Asn Asp Lys His Pro Gly
20 25 30

atg cat tgg ctg gat gag gaa aag act ctg gtt cgt tta cct tgg aat 144
Met His Trp Leu Asp Glu Glu Lys Thr Leu Val Arg Leu Pro Trp Asn
35 40 45

cat tta aag ggc gcg ggt ggc gtc tct gac gat gag aga aac atg tat 192
His Leu Lys Gly Ala Gly Gly Val Ser Asp Asp Glu Arg Asn Met Tyr
50 55 60

ctg gac tat tgt caa ttt aaa ggc ata cgg cag acc ggt aac aga cga 240
Leu Asp Tyr Cys Gln Phe Lys Gly Ile Arg Gln Thr Gly Asn Arg Arg
65 70 75 80

```
ttg agc gtc agg gaa tgt aag aac tgg ttg gcc agc gct ata cgc cac    288
Leu Ser Val Arg Glu Cys Lys Asn Trp Leu Ala Ser Ala Ile Arg His
                 85                  90                  95

agt cag acc gtc gaa gat gta tcc act gag gag aac ctg tcg gcg ccg    336
Ser Gln Thr Val Glu Asp Val Ser Thr Glu Glu Asn Leu Ser Ala Pro
            100                 105                 110

gcg cct aac agg tgc cgc gtt att cgc ctg ttg ccg atc ttt gta cga    384
Ala Pro Asn Arg Cys Arg Val Ile Arg Leu Leu Pro Ile Phe Val Arg
        115                 120                 125

tct tgc ccg ctc tgt aac gaa gcg gat gcc acc ggc ggc atg ctc ctg    432
Ser Cys Pro Leu Cys Asn Glu Ala Asp Ala Thr Gly Gly Met Leu Leu
    130                 135                 140

gac gta cgc aac gag gta acc gcc aga ttc cgg tat ctc ggt gcc ggg    480
Asp Val Arg Asn Glu Val Thr Ala Arg Phe Arg Tyr Leu Gly Ala Gly
145                 150                 155                 160

atg gag tac gaa ggc gcc gtg gga ggc gac ggg gag cag tgt tgg atg    528
Met Glu Tyr Glu Gly Ala Val Gly Gly Asp Gly Glu Gln Cys Trp Met
                165                 170                 175

cta cgg ctg gtt gtt tat tac tat ggg cgg ttg gtt ggg aac atg gag    576
Leu Arg Leu Val Val Tyr Tyr Tyr Gly Arg Leu Val Gly Asn Met Glu
            180                 185                 190

gtc ggg tcg ccc aac ggc gtt cgc ctg ttg ccg gcg cca aag cga cca    624
Val Gly Ser Pro Asn Gly Val Arg Leu Leu Pro Ala Pro Lys Arg Pro
        195                 200                 205

ctg caa gga cac gtt tgt gcg ggc att cgg ccg gaa cag gcc ctg ttg    672
Leu Gln Gly His Val Cys Ala Gly Ile Arg Pro Glu Gln Ala Leu Leu
    210                 215                 220

ccg cat acc cca cag gat atg ttt cct cac cag acg agc atg cta aag    720
Pro His Thr Pro Gln Asp Met Phe Pro His Gln Thr Ser Met Leu Lys
225                 230                 235                 240

tgg ctg ggg aag gag atc ata cgc ggg ttg atg att tac gca gac ggg    768
Trp Leu Gly Lys Glu Ile Ile Arg Gly Leu Met Ile Tyr Ala Asp Gly
                245                 250                 255

tct ggg att tac att cgg tat atg ggt cac gtt cca gcc ttc ctg ctg    816
Ser Gly Ile Tyr Ile Arg Tyr Met Gly His Val Pro Ala Phe Leu Leu
            260                 265                 270

ggt aac gga ggt tcg ctg gag ccg gtg gat ata att aac aac gcg cga    864
Gly Asn Gly Gly Ser Leu Glu Pro Val Asp Ile Ile Asn Asn Ala Arg
        275                 280                 285

gtt ctg cgc gtg ttt tcg ctc gcc caa tat ctg agt gcg gtg tca gcc    912
Val Leu Arg Val Phe Ser Leu Ala Gln Tyr Leu Ser Ala Val Ser Ala
    290                 295                 300

acc ccg cca cac gga acg cga ttt cca gcc gcc tat gcg tct ctc cac    960
Thr Pro Pro His Gly Thr Arg Phe Pro Ala Ala Tyr Ala Ser Leu His
305                 310                 315                 320

cta gga ggc gtt ccc act ccg gaa ggc gag ccg tgt ccc aca atc ccc   1008
```

Leu Gly Gly Val Pro Thr Pro Glu Gly Glu Pro Cys Pro Thr Ile Pro
325 330 335

ctg tcc att caa att tgg cac gag tgt ctg tgg cgg gcg tgc ggg gat 1056
Leu Ser Ile Gln Ile Trp His Glu Cys Leu Trp Arg Ala Cys Gly Asp
340 345 350

gcg gcc cag tga 1068
Ala Ala Gln
355

<210> 127
<211> 355
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 127
Met Ala Glu Gly Arg Ala Gly Ser Ile Arg Val Asn Arg Pro Ser Gly
1 5 10 15
Leu Arg Ala Trp Leu Leu Asp Cys Cys Asp Asn Asp Lys His Pro Gly
20 25 30
Met His Trp Leu Asp Glu Glu Lys Thr Leu Val Arg Leu Pro Trp Asn
35 40 45
His Leu Lys Gly Ala Gly Gly Val Ser Asp Asp Glu Arg Asn Met Tyr
50 55 60
Leu Asp Tyr Cys Gln Phe Lys Gly Ile Arg Gln Thr Gly Asn Arg Arg
65 70 75 80
Leu Ser Val Arg Glu Cys Lys Asn Trp Leu Ala Ser Ala Ile Arg His
85 90 95
Ser Gln Thr Val Glu Asp Val Ser Thr Glu Glu Asn Leu Ser Ala Pro
100 105 110
Ala Pro Asn Arg Cys Arg Val Ile Arg Leu Leu Pro Ile Phe Val Arg
115 120 125
Ser Cys Pro Leu Cys Asn Glu Ala Asp Ala Thr Gly Gly Met Leu Leu
130 135 140
Asp Val Arg Asn Glu Val Thr Ala Arg Phe Arg Tyr Leu Gly Ala Gly
145 150 155 160
Met Glu Tyr Glu Gly Ala Val Gly Gly Asp Gly Glu Gln Cys Trp Met
165 170 175
Leu Arg Leu Val Val Tyr Tyr Tyr Gly Arg Leu Val Gly Asn Met Glu
180 185 190
Val Gly Ser Pro Asn Gly Val Arg Leu Leu Pro Ala Pro Lys Arg Pro
195 200 205
Leu Gln Gly His Val Cys Ala Gly Ile Arg Pro Glu Gln Ala Leu Leu
210 215 220
Pro His Thr Pro Gln Asp Met Phe Pro His Gln Thr Ser Met Leu Lys
225 230 235 240
Trp Leu Gly Lys Glu Ile Ile Arg Gly Leu Met Ile Tyr Ala Asp Gly
245 250 255
Ser Gly Ile Tyr Ile Arg Tyr Met Gly His Val Pro Ala Phe Leu Leu
260 265 270
Gly Asn Gly Gly Ser Leu Glu Pro Val Asp Ile Ile Asn Asn Ala Arg
275 280 285
Val Leu Arg Val Phe Ser Leu Ala Gln Tyr Leu Ser Ala Val Ser Ala
290 295 300
Thr Pro Pro His Gly Thr Arg Phe Pro Ala Ala Tyr Ala Ser Leu His
305 310 315 320
Leu Gly Gly Val Pro Thr Pro Glu Gly Glu Pro Cys Pro Thr Ile Pro
325 330 335
Leu Ser Ile Gln Ile Trp His Glu Cys Leu Trp Arg Ala Cys Gly Asp

340 345 350
Ala Ala Gln
355

<210> 128
<211> 1095
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1095)

<400> 128
atg acc gag ata gaa atc act cac aat cac ctg cga aga tgg att att 48
Met Thr Glu Ile Glu Ile Thr His Asn His Leu Arg Arg Trp Ile Ile
1 5 10 15

tcc aat ttg gaa gcg aac acg ttt ccc gaa cac ttg tgt tgg tgc gac 96
Ser Asn Leu Glu Ala Asn Thr Phe Pro Glu His Leu Cys Trp Cys Asp
20 25 30

gag gaa aag agg agc ttt agg ata tca tgg cac cga ggt atg agt ggc 144
Glu Glu Lys Arg Ser Phe Arg Ile Ser Trp His Arg Gly Met Ser Gly
35 40 45

atg caa cct gtg gtt gcg tac tgc ctg gat agg gat ctg gaa tgt gga 192
Met Gln Pro Val Val Ala Tyr Cys Leu Asp Arg Asp Leu Glu Cys Gly
50 55 60

cgc cag cac aac gtg tcg gag tgc cgg aaa aga ctg cta agg gtg ctt 240
Arg Gln His Asn Val Ser Glu Cys Arg Lys Arg Leu Leu Arg Val Leu
65 70 75 80

cgg gaa aac gct ggg ttt gaa cag gac gat gcc cgg gca aca aca acg 288
Arg Glu Asn Ala Gly Phe Glu Gln Asp Asp Ala Arg Ala Thr Thr Thr
85 90 95

cgt ttt ggc gga gaa aga ttt ttt tac ttg aga ccc gcc gtg gat ccg 336
Arg Phe Gly Gly Glu Arg Phe Phe Tyr Leu Arg Pro Ala Val Asp Pro
100 105 110

ctg tgt tat gcc tgc atc tta gat agt cat tcg gag acc gtc cta aat 384
Leu Cys Tyr Ala Cys Ile Leu Asp Ser His Ser Glu Thr Val Leu Asn
115 120 125

tac tta gag gcg gcc tgc gtc cat ggg ctt gag cca ggg acg ccc ctg 432
Tyr Leu Glu Ala Ala Cys Val His Gly Leu Glu Pro Gly Thr Pro Leu
130 135 140

cct cct ccc gct ccg gcg gag gct gac ggt gcg gcg cgg tcc gta tat 480
Pro Pro Pro Ala Pro Ala Glu Ala Asp Gly Ala Ala Arg Ser Val Tyr
145 150 155 160

gcc agg gcc gcc agg ttg gcg acg gtt gca cca ccg cat cca gac cag 528
Ala Arg Ala Ala Arg Leu Ala Thr Val Ala Pro Pro His Pro Asp Gln
165 170 175

atc act cca ttt tgg cgg cta cgt ata cga gtt ttt tat ttt ggg tcg 576

Ile Thr Pro Phe Trp Arg Leu Arg Ile Arg Val Phe Tyr Phe Gly Ser
180 185 190

ctg gta gcg gaa cac acg tcc cag gac cgt aga ggg gta cga ttg cat 624
Leu Val Ala Glu His Thr Ser Gln Asp Arg Arg Gly Val Arg Leu His
195 200 205

aag cgc caa gac ccc aaa ccg ggt cac gag tgt ttc tat ggg aca gcg 672
Lys Arg Gln Asp Pro Lys Pro Gly His Glu Cys Phe Tyr Gly Thr Ala
210 215 220

tat aag atg tgg ctt cca aaa cct cag ttg gat ggg ccg cta acg ccg 720
Tyr Lys Met Trp Leu Pro Lys Pro Gln Leu Asp Gly Pro Leu Thr Pro
225 230 235 240

gag cag aga gaa acc gtg tgt gag att ata aac ggg tgt gag gag ggc 768
Glu Gln Arg Glu Thr Val Cys Glu Ile Ile Asn Gly Cys Glu Glu Gly
245 250 255

gtc ttt ttg cat ggc aat gag ctg ggg atg tat gtg gat aac aga acc 816
Val Phe Leu His Gly Asn Glu Leu Gly Met Tyr Val Asp Asn Arg Thr
260 265 270

agg cac acg gtt cgc tgc gca ggg aac gac gca gag ggg aac cac gca 864
Arg His Thr Val Arg Cys Ala Gly Asn Asp Ala Glu Gly Asn His Ala
275 280 285

caa cgg gct gtg cga tcc tct gtc aaa tct caa atc ttc tat gtt atg 912
Gln Arg Ala Val Arg Ser Ser Val Lys Ser Gln Ile Phe Tyr Val Met
290 295 300

ggt cta ctg cgc aga ctc gcc cgg tca ccc gtt ccc ggc gac act gtt 960
Gly Leu Leu Arg Arg Leu Ala Arg Ser Pro Val Pro Gly Asp Thr Val
305 310 315 320

ccc agc aac gca gtc act ctt tac ctt ggg ggt cgc cct ggc tcc agt 1008
Pro Ser Asn Ala Val Thr Leu Tyr Leu Gly Gly Arg Pro Gly Ser Ser
325 330 335

aaa aga ccc cag gtc cct gtc act ttg gtg atc tgt cag gat gaa ttg 1056
Lys Arg Pro Gln Val Pro Val Thr Leu Val Ile Cys Gln Asp Glu Leu
340 345 350

act cat ggt gac att cgg gcg gct cgg tgg att ttg tag 1095
Thr His Gly Asp Ile Arg Ala Ala Arg Trp Ile Leu
355 360 365

<210> 129
<211> 364
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 129
Met Thr Glu Ile Glu Ile Thr His Asn His Leu Arg Arg Trp Ile Ile
1 5 10 15
Ser Asn Leu Glu Ala Asn Thr Phe Pro Glu His Leu Cys Trp Cys Asp
20 25 30
Glu Glu Lys Arg Ser Phe Arg Ile Ser Trp His Arg Gly Met Ser Gly
35 40 45
Met Gln Pro Val Val Ala Tyr Cys Leu Asp Arg Asp Leu Glu Cys Gly

50 55 60
Arg Gln His Asn Val Ser Glu Cys Arg Lys Arg Leu Leu Arg Val Leu
65 70 75 80
Arg Glu Asn Ala Gly Phe Glu Gln Asp Asp Ala Arg Ala Thr Thr Thr
85 90 95
Arg Phe Gly Gly Glu Arg Phe Phe Tyr Leu Arg Pro Ala Val Asp Pro
100 105 110
Leu Cys Tyr Ala Cys Ile Leu Asp Ser His Ser Glu Thr Val Leu Asn
115 120 125
Tyr Leu Glu Ala Ala Cys Val His Gly Leu Glu Pro Gly Thr Pro Leu
130 135 140
Pro Pro Pro Ala Pro Ala Glu Ala Asp Gly Ala Ala Arg Ser Val Tyr
145 150 155 160
Ala Arg Ala Ala Arg Leu Ala Thr Val Ala Pro Pro His Pro Asp Gln
165 170 175
Ile Thr Pro Phe Trp Arg Leu Arg Ile Arg Val Phe Tyr Phe Gly Ser
180 185 190
Leu Val Ala Glu His Thr Ser Gln Asp Arg Arg Gly Val Arg Leu His
195 200 205
Lys Arg Gln Asp Pro Lys Pro Gly His Glu Cys Phe Tyr Gly Thr Ala
210 215 220
Tyr Lys Met Trp Leu Pro Lys Pro Gln Leu Asp Gly Pro Leu Thr Pro
225 230 235 240
Glu Gln Arg Glu Thr Val Cys Glu Ile Ile Asn Gly Cys Glu Glu Gly
245 250 255
Val Phe Leu His Gly Asn Glu Leu Gly Met Tyr Val Asp Asn Arg Thr
260 265 270
Arg His Thr Val Arg Cys Ala Gly Asn Asp Ala Glu Gly Asn His Ala
275 280 285
Gln Arg Ala Val Arg Ser Ser Val Lys Ser Gln Ile Phe Tyr Val Met
290 295 300
Gly Leu Leu Arg Arg Leu Ala Arg Ser Pro Val Pro Gly Asp Thr Val
305 310 315 320
Pro Ser Asn Ala Val Thr Leu Tyr Leu Gly Gly Arg Pro Gly Ser Ser
325 330 335
Lys Arg Pro Gln Val Pro Val Thr Leu Val Ile Cys Gln Asp Glu Leu
340 345 350
Thr His Gly Asp Ile Arg Ala Ala Arg Trp Ile Leu
355 360

<210> 130
<211> 1083
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1083)

<400> 130
atg ggc acg tac acg tcg gag gcg tcc ctc gcc tgg cta tct ttt atg 48
Met Gly Thr Tyr Thr Ser Glu Ala Ser Leu Ala Trp Leu Ser Phe Met
1 5 10 15

agc ggg acg gtc tcg gct tcc ccg ttt att ctg tgt ttt att tat cac 96
Ser Gly Thr Val Ser Ala Ser Pro Phe Ile Leu Cys Phe Ile Tyr His
20 25 30

tcg ctg tat ttt gta gag ccc ctg att agc gtt gag aac att ata ttc 144

Ser Leu Tyr Phe Val Glu Pro Leu Ile Ser Val Glu Asn Ile Ile Phe
35 40 45

tcc tgg ggc gcg gtt ggc tta cat ggg ctg ctt ttg cta ttc tgt ata 192
Ser Trp Gly Ala Val Gly Leu His Gly Leu Leu Leu Leu Phe Cys Ile
50 55 60

ttt ggg cta ccg gcc tgg ctc tcg cgg cag gtg gat gta ccg tgc acc 240
Phe Gly Leu Pro Ala Trp Leu Ser Arg Gln Val Asp Val Pro Cys Thr
65 70 75 80

atc tcg gcg ttt ctt att acc gcc ggt tct atg gca tcc acc ctc ggc 288
Ile Ser Ala Phe Leu Ile Thr Ala Gly Ser Met Ala Ser Thr Leu Gly
85 90 95

gtt gac ctt cca tgg gtt cac gtt tcc att ttt gtg ggg tcg tgc ctg 336
Val Asp Leu Pro Trp Val His Val Ser Ile Phe Val Gly Ser Cys Leu
100 105 110

tgc ctg ctg ctc tgt gtt gtt gcg gcg aat gac gtg gta tat ttg tgc 384
Cys Leu Leu Leu Cys Val Val Ala Ala Asn Asp Val Val Tyr Leu Cys
115 120 125

ccc acg att gcc cac aga tac tac gaa ctt gga ttt tta gca gcg ttc 432
Pro Thr Ile Ala His Arg Tyr Tyr Glu Leu Gly Phe Leu Ala Ala Phe
130 135 140

tcg gtg tat tat ttt tta gtt ctt aag aac ctt ttt tta gcc ccg gtg 480
Ser Val Tyr Tyr Phe Leu Val Leu Lys Asn Leu Phe Leu Ala Pro Val
145 150 155 160

ttt ttg ctg ccc ctg gtg gcc ttt ata gta ggt ggc gtt tgt tca ctc 528
Phe Leu Leu Pro Leu Val Ala Phe Ile Val Gly Gly Val Cys Ser Leu
165 170 175

aga gcg ctg cgg tca cat ccg ctt tat gag gcg ggc ctg cag cga cgc 576
Arg Ala Leu Arg Ser His Pro Leu Tyr Glu Ala Gly Leu Gln Arg Arg
180 185 190

cat gca att ttt tcg ctc acc tct agg cgg tac atc aca tac tct ata 624
His Ala Ile Phe Ser Leu Thr Ser Arg Arg Tyr Ile Thr Tyr Ser Ile
195 200 205

aag caa gcg ctt gag gtg tgt gga tgg gac ttt tat ctc gtg act gtt 672
Lys Gln Ala Leu Glu Val Cys Gly Trp Asp Phe Tyr Leu Val Thr Val
210 215 220

ttg att gga ggg gcg gcc gcg ggt acg ctg tct gtt ggc ctg acg acg 720
Leu Ile Gly Gly Ala Ala Ala Gly Thr Leu Ser Val Gly Leu Thr Thr
225 230 235 240

cct ctg ctc ctc gga ctt gta cac tat ttt ttt gtc ttc cac gtt ggg 768
Pro Leu Leu Leu Gly Leu Val His Tyr Phe Phe Val Phe His Val Gly
245 250 255

ctt ttt tgt tgc ctc gga ctc gtg ttt cgg tca aac gtt tta gcc ctg 816
Leu Phe Cys Cys Leu Gly Leu Val Phe Arg Ser Asn Val Leu Ala Leu
260 265 270

gtt tat gtt tta gca gcc gct gtg ctg ctg acg ctc acg cac gtc ttg 864
Val Tyr Val Leu Ala Ala Ala Val Leu Leu Thr Leu Thr His Val Leu

```
        275                 280                 285

ggg cca gga acg cat aat ttg ttc acc aga gtg tgt gtg ttt acg gtt     912
Gly Pro Gly Thr His Asn Leu Phe Thr Arg Val Cys Val Phe Thr Val
    290                 295                 300

ttt tta ttg act atg ttt ggg gcg att gga tgc gaa tta caa ata atc     960
Phe Leu Leu Thr Met Phe Gly Ala Ile Gly Cys Glu Leu Gln Ile Ile
305                 310                 315                 320

aga aaa aaa cta cag cgt gcc gcg aac tcg cca aga ata gtc ttg ggg    1008
Arg Lys Lys Leu Gln Arg Ala Ala Asn Ser Pro Arg Ile Val Leu Gly
                325                 330                 335

gtg tgt gcc tgc gga aac ctt ctg atg gcg gtg gtt ttt ttc tcc tta    1056
Val Cys Ala Cys Gly Asn Leu Leu Met Ala Val Val Phe Phe Ser Leu
            340                 345                 350

aat aaa gtt gag ctt ggt gcc ctt taa                                1083
Asn Lys Val Glu Leu Gly Ala Leu
        355                 360


<210> 131
<211> 360
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 131
Met Gly Thr Tyr Thr Ser Glu Ala Ser Leu Ala Trp Leu Ser Phe Met
 1               5                  10                  15
Ser Gly Thr Val Ser Ala Ser Pro Phe Ile Leu Cys Phe Ile Tyr His
            20                  25                  30
Ser Leu Tyr Phe Val Glu Pro Leu Ile Ser Val Glu Asn Ile Ile Phe
        35                  40                  45
Ser Trp Gly Ala Val Gly Leu His Gly Leu Leu Leu Leu Phe Cys Ile
    50                  55                  60
Phe Gly Leu Pro Ala Trp Leu Ser Arg Gln Val Asp Val Pro Cys Thr
65                  70                  75                  80
Ile Ser Ala Phe Leu Ile Thr Ala Gly Ser Met Ala Ser Thr Leu Gly
                85                  90                  95
Val Asp Leu Pro Trp Val His Val Ser Ile Phe Val Gly Ser Cys Leu
            100                 105                 110
Cys Leu Leu Leu Cys Val Val Ala Ala Asn Asp Val Val Tyr Leu Cys
        115                 120                 125
Pro Thr Ile Ala His Arg Tyr Tyr Glu Leu Gly Phe Leu Ala Ala Phe
    130                 135                 140
Ser Val Tyr Tyr Phe Leu Val Leu Lys Asn Leu Phe Leu Ala Pro Val
145                 150                 155                 160
Phe Leu Leu Pro Leu Val Ala Phe Ile Val Gly Gly Val Cys Ser Leu
                165                 170                 175
Arg Ala Leu Arg Ser His Pro Leu Tyr Glu Ala Gly Leu Gln Arg Arg
            180                 185                 190
His Ala Ile Phe Ser Leu Thr Ser Arg Arg Tyr Ile Thr Tyr Ser Ile
        195                 200                 205
Lys Gln Ala Leu Glu Val Cys Gly Trp Asp Phe Tyr Leu Val Thr Val
    210                 215                 220
Leu Ile Gly Gly Ala Ala Ala Gly Thr Leu Ser Val Gly Leu Thr Thr
225                 230                 235                 240
Pro Leu Leu Leu Gly Leu Val His Tyr Phe Phe Val Phe His Val Gly
                245                 250                 255
```

Leu Phe Cys Cys Leu Gly Leu Val Phe Arg Ser Asn Val Leu Ala Leu
260 265 270
Val Tyr Val Leu Ala Ala Ala Val Leu Leu Thr Leu Thr His Val Leu
275 280 285
Gly Pro Gly Thr His Asn Leu Phe Thr Arg Val Cys Val Phe Thr Val
290 295 300
Phe Leu Leu Thr Met Phe Gly Ala Ile Gly Cys Glu Leu Gln Ile Ile
305 310 315 320
Arg Lys Lys Leu Gln Arg Ala Ala Asn Ser Pro Arg Ile Val Leu Gly
325 330 335
Val Cys Ala Cys Gly Asn Leu Leu Met Ala Val Val Phe Phe Ser Leu
340 345 350
Asn Lys Val Glu Leu Gly Ala Leu
355 360

<210> 132
<211> 1185
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1185)

<400> 132
atg cct gtc tcg ttt cat tac ggc gcc agg gtg gac gtc gat gcg ctt 48
Met Pro Val Ser Phe His Tyr Gly Ala Arg Val Asp Val Asp Ala Leu
1 5 10 15

ggc agc att agc agg gtt tac gac cac att aag ggc atc gtt aag aaa 96
Gly Ser Ile Ser Arg Val Tyr Asp His Ile Lys Gly Ile Val Lys Lys
20 25 30

ggc gtt ata cag atc agc ggc cag ggc agg gcc ccc gtc ctt agc gtg 144
Gly Val Ile Gln Ile Ser Gly Gln Gly Arg Ala Pro Val Leu Ser Val
35 40 45

ctc tcc agc gtg ggg gac gcg ggc gtt ctc ggt ctc agg ctt aaa aat 192
Leu Ser Ser Val Gly Asp Ala Gly Val Leu Gly Leu Arg Leu Lys Asn
50 55 60

gca ctg gcg ccg ctg atg gta tat tct gac atg acg gac gag gtt agc 240
Ala Leu Ala Pro Leu Met Val Tyr Ser Asp Met Thr Asp Glu Val Ser
65 70 75 80

ttt agc ttt cga aac acc tcc ctt ggg aac acg ttc aca cac acc cgt 288
Phe Ser Phe Arg Asn Thr Ser Leu Gly Asn Thr Phe Thr His Thr Arg
85 90 95

gag atg ttt gga gtt aac att gcg gaa atg aac gtg gcg ttc tac cac 336
Glu Met Phe Gly Val Asn Ile Ala Glu Met Asn Val Ala Phe Tyr His
100 105 110

cac ggg gac gag tcg gac gcg gaa gga aag ccg cag ttt gtt agg acg 384
His Gly Asp Glu Ser Asp Ala Glu Gly Lys Pro Gln Phe Val Arg Thr
115 120 125

acc atc gcc tac ggg gac aat cac acg agc acc gtg cac aag agc gtc 432
Thr Ile Ala Tyr Gly Asp Asn His Thr Ser Thr Val His Lys Ser Val

130 135 140

gtg gac gaa ccc aac cta ccc tcg ttc cac gac agg cta gag cag gcc 480
Val Asp Glu Pro Asn Leu Pro Ser Phe His Asp Arg Leu Glu Gln Ala
145 150 155 160

ggg acg ggg aat cgc ctg ttc ctt acg gtt aag acg ctg act ttg ttg 528
Gly Thr Gly Asn Arg Leu Phe Leu Thr Val Lys Thr Leu Thr Leu Leu
165 170 175

tta aag tgg cta agg cag cag aaa acc agg gcc aag cag gtg gtc acc 576
Leu Lys Trp Leu Arg Gln Gln Lys Thr Arg Ala Lys Gln Val Val Thr
180 185 190

gtc tcg tta agc gaa acg ctg gca gtt gcc acg ttt acg gtc gac ggg 624
Val Ser Leu Ser Glu Thr Leu Ala Val Ala Thr Phe Thr Val Asp Gly
195 200 205

gtt agt aaa att att gat ttt aaa ccc gac acc ccg gac gcc aag tgg 672
Val Ser Lys Ile Ile Asp Phe Lys Pro Asp Thr Pro Asp Ala Lys Trp
210 215 220

act tgc gct aga gga aga aag ctg gac gtg ggt gtg gtt tct agc gac 720
Thr Cys Ala Arg Gly Arg Lys Leu Asp Val Gly Val Val Ser Ser Asp
225 230 235 240

ctg acg act cat gtg agt tta gag tct ctg gtg gcc gca ttg aac gca 768
Leu Thr Thr His Val Ser Leu Glu Ser Leu Val Ala Ala Leu Asn Ala
245 250 255

tgt aag att ccg gga ttt ttt ttg cct ggt ttc cgc tgg cac gcc aat 816
Cys Lys Ile Pro Gly Phe Phe Leu Pro Gly Phe Arg Trp His Ala Asn
260 265 270

gag att ctg gaa gtt gag ggc ctg ccg ctg aca gat agc ctt gct gac 864
Glu Ile Leu Glu Val Glu Gly Leu Pro Leu Thr Asp Ser Leu Ala Asp
275 280 285

gtc agg ttg ggc gtc atg cta tta aag gta gat ccg acg gac cgg aat 912
Val Arg Leu Gly Val Met Leu Leu Lys Val Asp Pro Thr Asp Arg Asn
290 295 300

aat gcc gtc ccg ggc aat ctc tcg gaa ggt gcc gat ccc gag ggg gtt 960
Asn Ala Val Pro Gly Asn Leu Ser Glu Gly Ala Asp Pro Glu Gly Val
305 310 315 320

cct gag ctc ccg tcc cct ccg aga acg cct gac ctg gac tta aaa gag 1008
Pro Glu Leu Pro Ser Pro Pro Arg Thr Pro Asp Leu Asp Leu Lys Glu
325 330 335

cag tgc gtc cca att gcc gag gac ggt gcc gag ccg aca gac ggg ggc 1056
Gln Cys Val Pro Ile Ala Glu Asp Gly Ala Glu Pro Thr Asp Gly Gly
340 345 350

gcc aaa tca ctg cgg acg agc ggc tct cga ccg gaa aaa aaa cac ggc 1104
Ala Lys Ser Leu Arg Thr Ser Gly Ser Arg Pro Glu Lys Lys His Gly
355 360 365

aag cga aaa cac agc agc tca ccc agt cgc ggg aaa gga aaa acc aag 1152
Lys Arg Lys His Ser Ser Ser Pro Ser Arg Gly Lys Gly Lys Thr Lys
370 375 380

acg cct cgc gcg acg ttc aac ccg ttg ttt taa 1185
Thr Pro Arg Ala Thr Phe Asn Pro Leu Phe
385 390 395

<210> 133
<211> 394
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 133
Met Pro Val Ser Phe His Tyr Gly Ala Arg Val Asp Val Asp Ala Leu
1 5 10 15
Gly Ser Ile Ser Arg Val Tyr Asp His Ile Lys Gly Ile Val Lys Lys
20 25 30
Gly Val Ile Gln Ile Ser Gly Gln Gly Arg Ala Pro Val Leu Ser Val
35 40 45
Leu Ser Ser Val Gly Asp Ala Gly Val Leu Gly Leu Arg Leu Lys Asn
50 55 60
Ala Leu Ala Pro Leu Met Val Tyr Ser Asp Met Thr Asp Glu Val Ser
65 70 75 80
Phe Ser Phe Arg Asn Thr Ser Leu Gly Asn Thr Phe Thr His Thr Arg
85 90 95
Glu Met Phe Gly Val Asn Ile Ala Glu Met Asn Val Ala Phe Tyr His
100 105 110
His Gly Asp Glu Ser Asp Ala Glu Gly Lys Pro Gln Phe Val Arg Thr
115 120 125
Thr Ile Ala Tyr Gly Asp Asn His Thr Ser Thr Val His Lys Ser Val
130 135 140
Val Asp Glu Pro Asn Leu Pro Ser Phe His Asp Arg Leu Glu Gln Ala
145 150 155 160
Gly Thr Gly Asn Arg Leu Phe Leu Thr Val Lys Thr Leu Thr Leu Leu
165 170 175
Leu Lys Trp Leu Arg Gln Gln Lys Thr Arg Ala Lys Gln Val Val Thr
180 185 190
Val Ser Leu Ser Glu Thr Leu Ala Val Ala Thr Phe Thr Val Asp Gly
195 200 205
Val Ser Lys Ile Ile Asp Phe Lys Pro Asp Thr Pro Asp Ala Lys Trp
210 215 220
Thr Cys Ala Arg Gly Arg Lys Leu Asp Val Gly Val Val Ser Ser Asp
225 230 235 240
Leu Thr Thr His Val Ser Leu Glu Ser Leu Val Ala Ala Leu Asn Ala
245 250 255
Cys Lys Ile Pro Gly Phe Phe Leu Pro Gly Phe Arg Trp His Ala Asn
260 265 270
Glu Ile Leu Glu Val Glu Gly Leu Pro Leu Thr Asp Ser Leu Ala Asp
275 280 285
Val Arg Leu Gly Val Met Leu Leu Lys Val Asp Pro Thr Asp Arg Asn
290 295 300
Asn Ala Val Pro Gly Asn Leu Ser Glu Gly Ala Asp Pro Glu Gly Val
305 310 315 320
Pro Glu Leu Pro Ser Pro Pro Arg Thr Pro Asp Leu Asp Leu Lys Glu
325 330 335
Gln Cys Val Pro Ile Ala Glu Asp Gly Ala Glu Pro Thr Asp Gly Gly
340 345 350
Ala Lys Ser Leu Arg Thr Ser Gly Ser Arg Pro Glu Lys Lys His Gly
355 360 365
Lys Arg Lys His Ser Ser Ser Pro Ser Arg Gly Lys Gly Lys Thr Lys
370 375 380
Thr Pro Arg Ala Thr Phe Asn Pro Leu Phe

385 390

<210> 134
<211> 945
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(945)

<400> 134
atg ttt ggc ctg tca ata gta acc gcc gcc atg gag tcc cca gat cgg 48
Met Phe Gly Leu Ser Ile Val Thr Ala Ala Met Glu Ser Pro Asp Arg
1 5 10 15

ttt ctg tac gcg agc gac cac cct ggc ttt tta gca ctc acc cag gag 96
Phe Leu Tyr Ala Ser Asp His Pro Gly Phe Leu Ala Leu Thr Gln Glu
20 25 30

aca tgg cag aac aga tgg ttc ccc agt caa att tcc ctg cac gag gac 144
Thr Trp Gln Asn Arg Trp Phe Pro Ser Gln Ile Ser Leu His Glu Asp
35 40 45

tcc gat gag gtg cga cta ttg agt ccc acc gat agg gag ttt tac cag 192
Ser Asp Glu Val Arg Leu Leu Ser Pro Thr Asp Arg Glu Phe Tyr Gln
50 55 60

ttt ctc ttc aca ttt ttg ggt atg gca gag agt ctg gta aac ttt aac 240
Phe Leu Phe Thr Phe Leu Gly Met Ala Glu Ser Leu Val Asn Phe Asn
65 70 75 80

atc gaa gac cta gtt aaa gag ttc agc aat cac gac gtg acc cac tat 288
Ile Glu Asp Leu Val Lys Glu Phe Ser Asn His Asp Val Thr His Tyr
85 90 95

tac gcc gaa cag gta gcc atg gaa aac atc cac ggc aaa gtg tat gcc 336
Tyr Ala Glu Gln Val Ala Met Glu Asn Ile His Gly Lys Val Tyr Ala
100 105 110

aac ata ctg aac tta ttt ttt ggc ggt aat cgg ggg gac tta atg att 384
Asn Ile Leu Asn Leu Phe Phe Gly Gly Asn Arg Gly Asp Leu Met Ile
115 120 125

tat gcc aaa aaa ata gtc gag gac gct acc ctg gcg aag aag att gac 432
Tyr Ala Lys Lys Ile Val Glu Asp Ala Thr Leu Ala Lys Lys Ile Asp
130 135 140

tgg ctg cat tcc cga gtg cgc aag gcc acc acg cgg gcc gag aag gtg 480
Trp Leu His Ser Arg Val Arg Lys Ala Thr Thr Arg Ala Glu Lys Val
145 150 155 160

ctg ctg ttc ctg gtt atc gaa ggg ata tac ttt ata agt tca ttt tac 528
Leu Leu Phe Leu Val Ile Glu Gly Ile Tyr Phe Ile Ser Ser Phe Tyr
165 170 175

tct ata gga ctt ttt cgc ctt cgc gga atc atg cgc ggg gtg tgc ctg 576
Ser Ile Gly Leu Phe Arg Leu Arg Gly Ile Met Arg Gly Val Cys Leu
180 185 190

gct aac gac tac atc tcc agg gac gag ctg cta cac acc agg gcc gcc 624
Ala Asn Asp Tyr Ile Ser Arg Asp Glu Leu Leu His Thr Arg Ala Ala
195 200 205

agc ttg ctg tac aac acc atg atc tcc cga gac gaa tcc ccc tcc gtc 672
Ser Leu Leu Tyr Asn Thr Met Ile Ser Arg Asp Glu Ser Pro Ser Val
210 215 220

gca tac atc cac ggc ctc ttt cgc gaa gcc gtg gag ata gag act ctg 720
Ala Tyr Ile His Gly Leu Phe Arg Glu Ala Val Glu Ile Glu Thr Leu
225 230 235 240

ttt att cgt tcc aag agc cga gac gtg acc atg gtg aac gtc gga gac 768
Phe Ile Arg Ser Lys Ser Arg Asp Val Thr Met Val Asn Val Gly Asp
245 250 255

atc gag cag ttt ttg cag gcc acg gcc gac agg att cta aag tca ata 816
Ile Glu Gln Phe Leu Gln Ala Thr Ala Asp Arg Ile Leu Lys Ser Ile
260 265 270

aac ata ccg ccc ctg ttc ggt gcc cgc ccc cca aac gcg tgt ccc ttg 864
Asn Ile Pro Pro Leu Phe Gly Ala Arg Pro Pro Asn Ala Cys Pro Leu
275 280 285

tcg tat acg agt gcc aag agt gtg aac ttt ttt gag cgt gat aat tct 912
Ser Tyr Thr Ser Ala Lys Ser Val Asn Phe Phe Glu Arg Asp Asn Ser
290 295 300

gaa tat gtc aca tcc gtt cat aat gat ctg tag 945
Glu Tyr Val Thr Ser Val His Asn Asp Leu
305 310 315

<210> 135
<211> 314
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 135
Met Phe Gly Leu Ser Ile Val Thr Ala Ala Met Glu Ser Pro Asp Arg
1 5 10 15
Phe Leu Tyr Ala Ser Asp His Pro Gly Phe Leu Ala Leu Thr Gln Glu
20 25 30
Thr Trp Gln Asn Arg Trp Phe Pro Ser Gln Ile Ser Leu His Glu Asp
35 40 45
Ser Asp Glu Val Arg Leu Leu Ser Pro Thr Asp Arg Glu Phe Tyr Gln
50 55 60
Phe Leu Phe Thr Phe Leu Gly Met Ala Glu Ser Leu Val Asn Phe Asn
65 70 75 80
Ile Glu Asp Leu Val Lys Glu Phe Ser Asn His Asp Val Thr His Tyr
85 90 95
Tyr Ala Glu Gln Val Ala Met Glu Asn Ile His Gly Lys Val Tyr Ala
100 105 110
Asn Ile Leu Asn Leu Phe Phe Gly Gly Asn Arg Gly Asp Leu Met Ile
115 120 125
Tyr Ala Lys Lys Ile Val Glu Asp Ala Thr Leu Ala Lys Lys Ile Asp
130 135 140
Trp Leu His Ser Arg Val Arg Lys Ala Thr Thr Arg Ala Glu Lys Val
145 150 155 160
Leu Leu Phe Leu Val Ile Glu Gly Ile Tyr Phe Ile Ser Ser Phe Tyr

```
                165                 170                 175
Ser Ile Gly Leu Phe Arg Leu Arg Gly Ile Met Arg Gly Val Cys Leu
            180                 185                 190
Ala Asn Asp Tyr Ile Ser Arg Asp Glu Leu Leu His Thr Arg Ala Ala
        195                 200                 205
Ser Leu Leu Tyr Asn Thr Met Ile Ser Arg Asp Glu Ser Pro Ser Val
    210                 215                 220
Ala Tyr Ile His Gly Leu Phe Arg Glu Ala Val Glu Ile Glu Thr Leu
225                 230                 235                 240
Phe Ile Arg Ser Lys Ser Arg Asp Val Thr Met Val Asn Val Gly Asp
                245                 250                 255
Ile Glu Gln Phe Leu Gln Ala Thr Ala Asp Arg Ile Leu Lys Ser Ile
            260                 265                 270
Asn Ile Pro Pro Leu Phe Gly Ala Arg Pro Pro Asn Ala Cys Pro Leu
        275                 280                 285
Ser Tyr Thr Ser Ala Lys Ser Val Asn Phe Phe Glu Arg Asp Asn Ser
    290                 295                 300
Glu Tyr Val Thr Ser Val His Asn Asp Leu
305                 310
```

```
<210> 136
<211> 2367
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2367)

<400> 136
atg aac acc gaa act tct ttc tct gcg gca aaa tct gcc aaa cct tta      48
Met Asn Thr Glu Thr Ser Phe Ser Ala Ala Lys Ser Ala Lys Pro Leu
  1               5                  10                  15

acc cta gtt acc gac gcg gga act ggg ggg tgt agc tca agc ctt gac      96
Thr Leu Val Thr Asp Ala Gly Thr Gly Gly Cys Ser Ser Ser Leu Asp
             20                  25                  30

cct gaa cga tgc gcc gaa tca ctc gtg aat tct tta aag gca aca ctt     144
Pro Glu Arg Cys Ala Glu Ser Leu Val Asn Ser Leu Lys Ala Thr Leu
         35                  40                  45

gga tgg gac ata gag gcc aac tca cta acc ggt ctc ctt tgg cac aga     192
Gly Trp Asp Ile Glu Ala Asn Ser Leu Thr Gly Leu Leu Trp His Arg
     50                  55                  60

ata atg gag gac agg tgc ctg gtt acg gtg cga gac tat ctc gcg gta     240
Ile Met Glu Asp Arg Cys Leu Val Thr Val Arg Asp Tyr Leu Ala Val
 65                  70                  75                  80

ttc ggg gaa agg cta agc gac gag gtt cgc gcg ttt atg tca aaa cac     288
Phe Gly Glu Arg Leu Ser Asp Glu Val Arg Ala Phe Met Ser Lys His
                 85                  90                  95

gaa gcg gcc ctt gac ggt ctt tta cag gac ttt aag cag agc aag gca     336
Glu Ala Ala Leu Asp Gly Leu Leu Gln Asp Phe Lys Gln Ser Lys Ala
            100                 105                 110

tac aca aat ttc gtt aac tgc ggc tac ctg tct gcc gtt agg ttt tac     384
```

Tyr Thr Asn Phe Val Asn Cys Gly Tyr Leu Ser Ala Val Arg Phe Tyr
115 120 125

gat acg tac gtg ctg aga acc cag ggt tct tca ccc att ttt gaa agc 432
Asp Thr Tyr Val Leu Arg Thr Gln Gly Ser Ser Pro Ile Phe Glu Ser
130 135 140

gtg gcc cag atg ttt atg agg gtg gct gtt ttt gtg gcc tgt cag tgt 480
Val Ala Gln Met Phe Met Arg Val Ala Val Phe Val Ala Cys Gln Cys
145 150 155 160

ata aaa ttt ccg tgt ttg cgg gag acc ctc aga cac ctg gtg gag agc 528
Ile Lys Phe Pro Cys Leu Arg Glu Thr Leu Arg His Leu Val Glu Ser
165 170 175

gaa acg gag ctg gat gag atg tac ctg gtt ggt tac gcg ttc cat tac 576
Glu Thr Glu Leu Asp Glu Met Tyr Leu Val Gly Tyr Ala Phe His Tyr
180 185 190

atc tcg tca cag att gtg tgc tgc gca acg cct gta tta agg tca gcc 624
Ile Ser Ser Gln Ile Val Cys Cys Ala Thr Pro Val Leu Arg Ser Ala
195 200 205

ggg ctt cgc gga ggt cag ctt tct agc tgt ttt att tta aag ccg tcc 672
Gly Leu Arg Gly Gly Gln Leu Ser Ser Cys Phe Ile Leu Lys Pro Ser
210 215 220

atg gcg acc gaa gac aaa acc ctc aaa gca ctt cat gag gaa atg tcc 720
Met Ala Thr Glu Asp Lys Thr Leu Lys Ala Leu His Glu Glu Met Ser
225 230 235 240

ccc ctt ctc gct agc aag tca ggc gtg ggg ata gat gtc tcg tct ttt 768
Pro Leu Leu Ala Ser Lys Ser Gly Val Gly Ile Asp Val Ser Ser Phe
245 250 255

gcg gaa cac aag aac atc acc agc tgt tta aaa tta ata aac gcc cac 816
Ala Glu His Lys Asn Ile Thr Ser Cys Leu Lys Leu Ile Asn Ala His
260 265 270

gtt ggg tat ttc aac gat aac aat atc cgt cct gtt ggt gcc agc gcc 864
Val Gly Tyr Phe Asn Asp Asn Asn Ile Arg Pro Val Gly Ala Ser Ala
275 280 285

tac atg gaa cta tgg cat cat cag att tgt gat ttt ttg aac gcg aag 912
Tyr Met Glu Leu Trp His His Gln Ile Cys Asp Phe Leu Asn Ala Lys
290 295 300

atg ccg gag aat cag gag cgg tgt cac aac ctg ttt cag ggc gtg tgt 960
Met Pro Glu Asn Gln Glu Arg Cys His Asn Leu Phe Gln Gly Val Cys
305 310 315 320

gtg ccg gag ctt ttt ttc cga ctg tac gaa aca aac cca gat ggc cag 1008
Val Pro Glu Leu Phe Phe Arg Leu Tyr Glu Thr Asn Pro Asp Gly Gln
325 330 335

tgg cac ctg ttt gcg cca gag gtg gcg cca aat ctc cta aaa ctc tac 1056
Trp His Leu Phe Ala Pro Glu Val Ala Pro Asn Leu Leu Lys Leu Tyr
340 345 350

ggt gcc gaa ttt gag att gaa tac aac aga cta gtg gcc gct ggc aaa 1104
Gly Ala Glu Phe Glu Ile Glu Tyr Asn Arg Leu Val Ala Ala Gly Lys

355 360 365

cac tct agc tcg ttg ccg tta aaa tcg atg atg tat gcc ctt att aac 1152
His Ser Ser Ser Leu Pro Leu Lys Ser Met Met Tyr Ala Leu Ile Asn
370 375 380

act gtc att aag acc ggc agt ccc tac gtc ctg cta aag gag gcg ctc 1200
Thr Val Ile Lys Thr Gly Ser Pro Tyr Val Leu Leu Lys Glu Ala Leu
385 390 395 400

aac aaa cac cac tgg tgt gag acg caa ggc agc gcc ata aac tgt tct 1248
Asn Lys His His Trp Cys Glu Thr Gln Gly Ser Ala Ile Asn Cys Ser
405 410 415

aat ctc tgt gcg gaa atc gtg cag caa ccc gag ggt cag gct tcc gtg 1296
Asn Leu Cys Ala Glu Ile Val Gln Gln Pro Glu Gly Gln Ala Ser Val
420 425 430

tgt aac ttg gca aac att agc cta cca aag tgc ctc aga ccc cac agg 1344
Cys Asn Leu Ala Asn Ile Ser Leu Pro Lys Cys Leu Arg Pro His Arg
435 440 445

gga gag tct ggt gtg gag cct ggg aag ggg gat gta aca ttt ggt ttt 1392
Gly Glu Ser Gly Val Glu Pro Gly Lys Gly Asp Val Thr Phe Gly Phe
450 455 460

gaa ctg ctc gat gat gcc gtg gag gcc gca gtt atc att gtc aat gcg 1440
Glu Leu Leu Asp Asp Ala Val Glu Ala Ala Val Ile Ile Val Asn Ala
465 470 475 480

tgc att ttg ggc ggg acg gcg cct acc gag agc gtt agg cgc ggc cag 1488
Cys Ile Leu Gly Gly Thr Ala Pro Thr Glu Ser Val Arg Arg Gly Gln
485 490 495

gag gaa cga tca atg ggc att ggc gtc caa gga ctg gcc gac gtg ttt 1536
Glu Glu Arg Ser Met Gly Ile Gly Val Gln Gly Leu Ala Asp Val Phe
500 505 510

gcc gaa ctg ggc ttt ggc tac ctt gac gcc gag agc gcc aag ctg gac 1584
Ala Glu Leu Gly Phe Gly Tyr Leu Asp Ala Glu Ser Ala Lys Leu Asp
515 520 525

gtg gag att ttt cag gca atg tat ttt acg gcg gtg caa acc agt cat 1632
Val Glu Ile Phe Gln Ala Met Tyr Phe Thr Ala Val Gln Thr Ser His
530 535 540

gag att gtt tta ctc ggc gag ggg act ccg ttt cgt gga tgg gag cgg 1680
Glu Ile Val Leu Leu Gly Glu Gly Thr Pro Phe Arg Gly Trp Glu Arg
545 550 555 560

agt cga ttg gct cag ggt gtg ttt cac tgg caa acg tgg gac ggc gtt 1728
Ser Arg Leu Ala Gln Gly Val Phe His Trp Gln Thr Trp Asp Gly Val
565 570 575

aag ccg tcc cac cct ccc ttg gag aga tgg gag cag ctg ggg cgc agc 1776
Lys Pro Ser His Pro Pro Leu Glu Arg Trp Glu Gln Leu Gly Arg Ser
580 585 590

att gcc cag cat gga att ttc aac agc cag ttt ttg gcg ctg atg cca 1824
Ile Ala Gln His Gly Ile Phe Asn Ser Gln Phe Leu Ala Leu Met Pro
595 600 605

acg gcc gga acc tct cag ctt acg ggg tac acc gaa gca ttt tat ccg 1872
Thr Ala Gly Thr Ser Gln Leu Thr Gly Tyr Thr Glu Ala Phe Tyr Pro
610 615 620

ttc ttt gcc aat atc gca tcg aag gtc acc agc aaa gaa gag atc ttg 1920
Phe Phe Ala Asn Ile Ala Ser Lys Val Thr Ser Lys Glu Glu Ile Leu
625 630 635 640

aag cca aac gtg act ttt ttt aag aga gtg aag cct ggt gac ctg cga 1968
Lys Pro Asn Val Thr Phe Phe Lys Arg Val Lys Pro Gly Asp Leu Arg
645 650 655

acg gtg cgc cgt tac gga ggt gac gtg gct tca ttt cca gaa ccc ctg 2016
Thr Val Arg Arg Tyr Gly Gly Asp Val Ala Ser Phe Pro Glu Pro Leu
660 665 670

aag gac aga tac aag att ttt ttg acg gcc ttt gat tac tgc cca att 2064
Lys Asp Arg Tyr Lys Ile Phe Leu Thr Ala Phe Asp Tyr Cys Pro Ile
675 680 685

aag cag ttg gag cgg gcc ggt gcg cgg gct cca ttt gta gac cag agt 2112
Lys Gln Leu Glu Arg Ala Gly Ala Arg Ala Pro Phe Val Asp Gln Ser
690 695 700

cag tcg ctt aat ttc ttt tta aag gag gag cag gcc acg cgg gcg agt 2160
Gln Ser Leu Asn Phe Phe Leu Lys Glu Glu Gln Ala Thr Arg Ala Ser
705 710 715 720

tat att cga gac ttg ctc ctg acc gga tac agg cta ggg tta aaa acc 2208
Tyr Ile Arg Asp Leu Leu Leu Thr Gly Tyr Arg Leu Gly Leu Lys Thr
725 730 735

atg ctg tat tat tgt cga att cag aaa cag aca aaa ctt aat gcc ctg 2256
Met Leu Tyr Tyr Cys Arg Ile Gln Lys Gln Thr Lys Leu Asn Ala Leu
740 745 750

cag tgc tta gat cag gtc gtg ggt gac aac atc tct tca gag ggg gcg 2304
Gln Cys Leu Asp Gln Val Val Gly Asp Asn Ile Ser Ser Glu Gly Ala
755 760 765

gag tct aac tgc gta caa aag gcc gac ggt gag cgg acc aag gta tgt 2352
Glu Ser Asn Cys Val Gln Lys Ala Asp Gly Glu Arg Thr Lys Val Cys
770 775 780

ttg gcc tgt caa tag 2367
Leu Ala Cys Gln
785

<210> 137
<211> 788
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 137
Met Asn Thr Glu Thr Ser Phe Ser Ala Ala Lys Ser Ala Lys Pro Leu
1 5 10 15
Thr Leu Val Thr Asp Ala Gly Thr Gly Gly Cys Ser Ser Ser Leu Asp
20 25 30
Pro Glu Arg Cys Ala Glu Ser Leu Val Asn Ser Leu Lys Ala Thr Leu

35 40 45
Gly Trp Asp Ile Glu Ala Asn Ser Leu Thr Gly Leu Leu Trp His Arg
50 55 60
Ile Met Glu Asp Arg Cys Leu Val Thr Val Arg Asp Tyr Leu Ala Val
65 70 75 80
Phe Gly Glu Arg Leu Ser Asp Glu Val Arg Ala Phe Met Ser Lys His
85 90 95
Glu Ala Ala Leu Asp Gly Leu Leu Gln Asp Phe Lys Gln Ser Lys Ala
100 105 110
Tyr Thr Asn Phe Val Asn Cys Gly Tyr Leu Ser Ala Val Arg Phe Tyr
115 120 125
Asp Thr Tyr Val Leu Arg Thr Gln Gly Ser Ser Pro Ile Phe Glu Ser
130 135 140
Val Ala Gln Met Phe Met Arg Val Ala Val Phe Val Ala Cys Gln Cys
145 150 155 160
Ile Lys Phe Pro Cys Leu Arg Glu Thr Leu Arg His Leu Val Glu Ser
165 170 175
Glu Thr Glu Leu Asp Glu Met Tyr Leu Val Gly Tyr Ala Phe His Tyr
180 185 190
Ile Ser Ser Gln Ile Val Cys Cys Ala Thr Pro Val Leu Arg Ser Ala
195 200 205
Gly Leu Arg Gly Gly Gln Leu Ser Ser Cys Phe Ile Leu Lys Pro Ser
210 215 220
Met Ala Thr Glu Asp Lys Thr Leu Lys Ala Leu His Glu Glu Met Ser
225 230 235 240
Pro Leu Leu Ala Ser Lys Ser Gly Val Gly Ile Asp Val Ser Ser Phe
245 250 255
Ala Glu His Lys Asn Ile Thr Ser Cys Leu Lys Leu Ile Asn Ala His
260 265 270
Val Gly Tyr Phe Asn Asp Asn Asn Ile Arg Pro Val Gly Ala Ser Ala
275 280 285
Tyr Met Glu Leu Trp His His Gln Ile Cys Asp Phe Leu Asn Ala Lys
290 295 300
Met Pro Glu Asn Gln Glu Arg Cys His Asn Leu Phe Gln Gly Val Cys
305 310 315 320
Val Pro Glu Leu Phe Phe Arg Leu Tyr Glu Thr Asn Pro Asp Gly Gln
325 330 335
Trp His Leu Phe Ala Pro Glu Val Ala Pro Asn Leu Leu Lys Leu Tyr
340 345 350
Gly Ala Glu Phe Glu Ile Glu Tyr Asn Arg Leu Val Ala Ala Gly Lys
355 360 365
His Ser Ser Ser Leu Pro Leu Lys Ser Met Met Tyr Ala Leu Ile Asn
370 375 380
Thr Val Ile Lys Thr Gly Ser Pro Tyr Val Leu Leu Lys Glu Ala Leu
385 390 395 400
Asn Lys His His Trp Cys Glu Thr Gln Gly Ser Ala Ile Asn Cys Ser
405 410 415
Asn Leu Cys Ala Glu Ile Val Gln Gln Pro Glu Gly Gln Ala Ser Val
420 425 430
Cys Asn Leu Ala Asn Ile Ser Leu Pro Lys Cys Leu Arg Pro His Arg
435 440 445
Gly Glu Ser Gly Val Glu Pro Gly Lys Gly Asp Val Thr Phe Gly Phe
450 455 460
Glu Leu Leu Asp Asp Ala Val Glu Ala Ala Val Ile Ile Val Asn Ala
465 470 475 480
Cys Ile Leu Gly Gly Thr Ala Pro Thr Glu Ser Val Arg Arg Gly Gln
485 490 495
Glu Glu Arg Ser Met Gly Ile Gly Val Gln Gly Leu Ala Asp Val Phe
500 505 510
Ala Glu Leu Gly Phe Gly Tyr Leu Asp Ala Glu Ser Ala Lys Leu Asp
515 520 525

Val Glu Ile Phe Gln Ala Met Tyr Phe Thr Ala Val Gln Thr Ser His
530 535 540
Glu Ile Val Leu Leu Gly Glu Gly Thr Pro Phe Arg Gly Trp Glu Arg
545 550 555 560
Ser Arg Leu Ala Gln Gly Val Phe His Trp Gln Thr Trp Asp Gly Val
565 570 575
Lys Pro Ser His Pro Pro Leu Glu Arg Trp Glu Gln Leu Gly Arg Ser
580 585 590
Ile Ala Gln His Gly Ile Phe Asn Ser Gln Phe Leu Ala Leu Met Pro
595 600 605
Thr Ala Gly Thr Ser Gln Leu Thr Gly Tyr Thr Glu Ala Phe Tyr Pro
610 615 620
Phe Phe Ala Asn Ile Ala Ser Lys Val Thr Ser Lys Glu Glu Ile Leu
625 630 635 640
Lys Pro Asn Val Thr Phe Phe Lys Arg Val Lys Pro Gly Asp Leu Arg
645 650 655
Thr Val Arg Arg Tyr Gly Gly Asp Val Ala Ser Phe Pro Glu Pro Leu
660 665 670
Lys Asp Arg Tyr Lys Ile Phe Leu Thr Ala Phe Asp Tyr Cys Pro Ile
675 680 685
Lys Gln Leu Glu Arg Ala Gly Ala Arg Ala Pro Phe Val Asp Gln Ser
690 695 700
Gln Ser Leu Asn Phe Phe Leu Lys Glu Glu Gln Ala Thr Arg Ala Ser
705 710 715 720
Tyr Ile Arg Asp Leu Leu Leu Thr Gly Tyr Arg Leu Gly Leu Lys Thr
725 730 735
Met Leu Tyr Tyr Cys Arg Ile Gln Lys Gln Thr Lys Leu Asn Ala Leu
740 745 750
Gln Cys Leu Asp Gln Val Val Gly Asp Asn Ile Ser Ser Glu Gly Ala
755 760 765
Glu Ser Asn Cys Val Gln Lys Ala Asp Gly Glu Arg Thr Lys Val Cys
770 775 780
Leu Ala Cys Gln
785

<210> 138
<211> 996
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(996)

<400> 138
atg aag acc agg gac gca aac gtg aac aag ctg aat gat agt ttg atg 48
Met Lys Thr Arg Asp Ala Asn Val Asn Lys Leu Asn Asp Ser Leu Met
1 5 10 15

cgt cta ctg ccg ccg cct cct cac cgg gtt tcg tta tct cgc ggt cgc 96
Arg Leu Leu Pro Pro Pro Pro His Arg Val Ser Leu Ser Arg Gly Arg
20 25 30

gac ttt tct aag ggt gtt cgt gat ctg ctc tct aaa tat gtg gtg tcg 144
Asp Phe Ser Lys Gly Val Arg Asp Leu Leu Ser Lys Tyr Val Val Ser
35 40 45

acc acc act ggg gtg gag gcc atc aaa gat gga ttt tta tct gtc tcg 192
Thr Thr Thr Gly Val Glu Ala Ile Lys Asp Gly Phe Leu Ser Val Ser

50 55 60

cct aca tgt caa acg tat gga gat ttt ttg att tat tct cag acc atg 240
Pro Thr Cys Gln Thr Tyr Gly Asp Phe Leu Ile Tyr Ser Gln Thr Met
65 70 75 80

agt tca cag gaa cct cgg ggc acc tac ctg ttt tcc ttt aaa cag act 288
Ser Ser Gln Glu Pro Arg Gly Thr Tyr Leu Phe Ser Phe Lys Gln Thr
85 90 95

gac act ggg tcc tcc att gac atg tta ttc acg cct acg tcg ctg gcg 336
Asp Thr Gly Ser Ser Ile Asp Met Leu Phe Thr Pro Thr Ser Leu Ala
100 105 110

cga ctg tct agg atg gac gcg gac tct gcc ccc cag acc aac aga ata 384
Arg Leu Ser Arg Met Asp Ala Asp Ser Ala Pro Gln Thr Asn Arg Ile
115 120 125

gcg tgt gtt tgg tac ggc cac gag agc ggt cta ctt gac gcc ata ccc 432
Ala Cys Val Trp Tyr Gly His Glu Ser Gly Leu Leu Asp Ala Ile Pro
130 135 140

aat ttt gaa gaa ttg ttg gaa acc ggc tcg ctt cat cag ttt ctt gcg 480
Asn Phe Glu Glu Leu Leu Glu Thr Gly Ser Leu His Gln Phe Leu Ala
145 150 155 160

ccg gtc ggg ccg ttg gtt caa acg gtc cac agt acg ttt gtc acc aag 528
Pro Val Gly Pro Leu Val Gln Thr Val His Ser Thr Phe Val Thr Lys
165 170 175

gtg acg tct gcg tta aag ggc aac gta gtc gcc aga gag ccc gtt gtt 576
Val Thr Ser Ala Leu Lys Gly Asn Val Val Ala Arg Glu Pro Val Val
180 185 190

aca cat att ggc ctt acg ctt cct agt gac atg ttt gtg gat tta gat 624
Thr His Ile Gly Leu Thr Leu Pro Ser Asp Met Phe Val Asp Leu Asp
195 200 205

gac tca tgc ccg tcc agc ctg cgc gac gag cct ttg ccc gcc cac tcg 672
Asp Ser Cys Pro Ser Ser Leu Arg Asp Glu Pro Leu Pro Ala His Ser
210 215 220

tct att tat gtt tgc tta acc tac att agg gta aat aat aga ccg gcg 720
Ser Ile Tyr Val Cys Leu Thr Tyr Ile Arg Val Asn Asn Arg Pro Ala
225 230 235 240

cta gga ctt ggg ttt ttt aag tct ggt aaa gga tat tgt gaa ata gct 768
Leu Gly Leu Gly Phe Phe Lys Ser Gly Lys Gly Tyr Cys Glu Ile Ala
245 250 255

gcc caa ttg agg gat ttt tac agc ggc gtg att cgc acc aaa tac att 816
Ala Gln Leu Arg Asp Phe Tyr Ser Gly Val Ile Arg Thr Lys Tyr Ile
260 265 270

cag ctg cag aat gac ctg tat ata aat agg ctt gcg ttc ggc gta gtt 864
Gln Leu Gln Asn Asp Leu Tyr Ile Asn Arg Leu Ala Phe Gly Val Val
275 280 285

tgc aga ctt gga agc gtg cct tcc ggt tta caa cca tct ttt caa agc 912
Cys Arg Leu Gly Ser Val Pro Ser Gly Leu Gln Pro Ser Phe Gln Ser
290 295 300

ctt cac ttt aaa ggc gct gct tta ccg gtg tta aaa ttt acg gag ttt 960
Leu His Phe Lys Gly Ala Ala Leu Pro Val Leu Lys Phe Thr Glu Phe
305 310 315 320

gtt tcc aat ccg ggg tct tgg aag cta ttt ctt taa 996
Val Ser Asn Pro Gly Ser Trp Lys Leu Phe Leu
325 330

<210> 139
<211> 331
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 139
Met Lys Thr Arg Asp Ala Asn Val Asn Lys Leu Asn Asp Ser Leu Met
1 5 10 15
Arg Leu Leu Pro Pro Pro Pro His Arg Val Ser Leu Ser Arg Gly Arg
20 25 30
Asp Phe Ser Lys Gly Val Arg Asp Leu Leu Ser Lys Tyr Val Val Ser
35 40 45
Thr Thr Thr Gly Val Glu Ala Ile Lys Asp Gly Phe Leu Ser Val Ser
50 55 60
Pro Thr Cys Gln Thr Tyr Gly Asp Phe Leu Ile Tyr Ser Gln Thr Met
65 70 75 80
Ser Ser Gln Glu Pro Arg Gly Thr Tyr Leu Phe Ser Phe Lys Gln Thr
85 90 95
Asp Thr Gly Ser Ser Ile Asp Met Leu Phe Thr Pro Thr Ser Leu Ala
100 105 110
Arg Leu Ser Arg Met Asp Ala Asp Ser Ala Pro Gln Thr Asn Arg Ile
115 120 125
Ala Cys Val Trp Tyr Gly His Glu Ser Gly Leu Leu Asp Ala Ile Pro
130 135 140
Asn Phe Glu Glu Leu Leu Glu Thr Gly Ser Leu His Gln Phe Leu Ala
145 150 155 160
Pro Val Gly Pro Leu Val Gln Thr Val His Ser Thr Phe Val Thr Lys
165 170 175
Val Thr Ser Ala Leu Lys Gly Asn Val Val Ala Arg Glu Pro Val Val
180 185 190
Thr His Ile Gly Leu Thr Leu Pro Ser Asp Met Phe Val Asp Leu Asp
195 200 205
Asp Ser Cys Pro Ser Ser Leu Arg Asp Glu Pro Leu Pro Ala His Ser
210 215 220
Ser Ile Tyr Val Cys Leu Thr Tyr Ile Arg Val Asn Asn Arg Pro Ala
225 230 235 240
Leu Gly Leu Gly Phe Phe Lys Ser Gly Lys Gly Tyr Cys Glu Ile Ala
245 250 255
Ala Gln Leu Arg Asp Phe Tyr Ser Gly Val Ile Arg Thr Lys Tyr Ile
260 265 270
Gln Leu Gln Asn Asp Leu Tyr Ile Asn Arg Leu Ala Phe Gly Val Val
275 280 285
Cys Arg Leu Gly Ser Val Pro Ser Gly Leu Gln Pro Ser Phe Gln Ser
290 295 300
Leu His Phe Lys Gly Ala Ala Leu Pro Val Leu Lys Phe Thr Glu Phe
305 310 315 320
Val Ser Asn Pro Gly Ser Trp Lys Leu Phe Leu
325 330

```
<210> 140
<211> 2820
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(2820)

<400> 140
atg gcg tcg tca att cct gcc gca cgg gct gat aac ggt gac gag aac      48
Met Ala Ser Ser Ile Pro Ala Ala Arg Ala Asp Asn Gly Asp Glu Asn
1               5                   10                  15

acc ggc ggg ctt tac aaa tta act gac aac ctc cta acg tgc acc gga      96
Thr Gly Gly Leu Tyr Lys Leu Thr Asp Asn Leu Leu Thr Cys Thr Gly
            20                  25                  30

tcg cta caa cag ctt aaa ctc ctg atg gag ttc caa cta aaa cca cta     144
Ser Leu Gln Gln Leu Lys Leu Leu Met Glu Phe Gln Leu Lys Pro Leu
        35                  40                  45

cca acc gca cac ctt tta agc atg ccc acc gtg acc cgg ttt tta aat     192
Pro Thr Ala His Leu Leu Ser Met Pro Thr Val Thr Arg Phe Leu Asn
    50                  55                  60

act gca ttc aaa ata gac aac ccc ctg gtt tca ttt att caa aaa cac     240
Thr Ala Phe Lys Ile Asp Asn Pro Leu Val Ser Phe Ile Gln Lys His
65                  70                  75                  80

cct gtg ttt ttt tta atg aga gtc gcc agg ctt ccg gag cca gtc att     288
Pro Val Phe Phe Leu Met Arg Val Ala Arg Leu Pro Glu Pro Val Ile
                85                  90                  95

aca gac cac caa agc gca gaa acg tca aca ggc ata cta tcc gag gtt     336
Thr Asp His Gln Ser Ala Glu Thr Ser Thr Gly Ile Leu Ser Glu Val
            100                 105                 110

gtg aat gtt ctt aat aca gct att cga aaa cca cac gag tcc cca gcg     384
Val Asn Val Leu Asn Thr Ala Ile Arg Lys Pro His Glu Ser Pro Ala
        115                 120                 125

gct aaa gac aac gat tat ctc gac aac cgt gcc ata ttg gcc atg att     432
Ala Lys Asp Asn Asp Tyr Leu Asp Asn Arg Ala Ile Leu Ala Met Ile
    130                 135                 140

aca gaa tac atc cat cac gta act tca cgt acg ccc tcg ggg atc cca     480
Thr Glu Tyr Ile His His Val Thr Ser Arg Thr Pro Ser Gly Ile Pro
145                 150                 155                 160

ccg aca cca cca atg ggt atc agc cat cta ccg tgc gta gag caa att     528
Pro Thr Pro Pro Met Gly Ile Ser His Leu Pro Cys Val Glu Gln Ile
                165                 170                 175

tta cac gaa acc cac cgg caa tac tgg aac cta acc ctc ccg gag tcg     576
Leu His Glu Thr His Arg Gln Tyr Trp Asn Leu Thr Leu Pro Glu Ser
            180                 185                 190

cta ttt atc gac atc ggc gaa gtc gcg tct ccg ctt cag acg tgg cta     624
Leu Phe Ile Asp Ile Gly Glu Val Ala Ser Pro Leu Gln Thr Trp Leu
        195                 200                 205
```

```
atc cta tcg tat tgt aaa aaa ctg caa ctg gca cca ccg ccc cta ttc     672
Ile Leu Ser Tyr Cys Lys Lys Leu Gln Leu Ala Pro Pro Pro Leu Phe
    210                 215                 220

cca cct gtc gac gaa ctg gcc cgt cgt ctg gtt acc ggc cac cac gaa     720
Pro Pro Val Asp Glu Leu Ala Arg Arg Leu Val Thr Gly His His Glu
225                 230                 235                 240

ttg ttt gtc ccc cta tcg acg tca ctg gaa acg tac atc acc atg cca     768
Leu Phe Val Pro Leu Ser Thr Ser Leu Glu Thr Tyr Ile Thr Met Pro
                245                 250                 255

gta tca aaa cag cga gcg ttt gag ata tac agc gtc ttt gcc aaa tca     816
Val Ser Lys Gln Arg Ala Phe Glu Ile Tyr Ser Val Phe Ala Lys Ser
            260                 265                 270

aaa aac ata gtt gac ggc aca cca att tta gcg ttc acc gac aca gaa     864
Lys Asn Ile Val Asp Gly Thr Pro Ile Leu Ala Phe Thr Asp Thr Glu
        275                 280                 285

ctt aca acg ttc acc ccc gaa ctg tta ttc ctg tac gac ttc gta atc     912
Leu Thr Thr Phe Thr Pro Glu Leu Leu Phe Leu Tyr Asp Phe Val Ile
    290                 295                 300

gag tcc ctg tgc aaa aac cag gca tac ggg tgt tcg cgc aac gca att     960
Glu Ser Leu Cys Lys Asn Gln Ala Tyr Gly Cys Ser Arg Asn Ala Ile
305                 310                 315                 320

gag cat ttt att aag aaa ggt atc gat ttt atg gcg gag ttg ggg gcg    1008
Glu His Phe Ile Lys Lys Gly Ile Asp Phe Met Ala Glu Leu Gly Ala
                325                 330                 335

ttc att gaa aaa acg tgt ggc tat cga tca acc gta agc ctg tcc aac    1056
Phe Ile Glu Lys Thr Cys Gly Tyr Arg Ser Thr Val Ser Leu Ser Asn
            340                 345                 350

gtg aga gcc gtt aag gcc agg ctc gcg agc tgt ggc cta tcc aaa gag    1104
Val Arg Ala Val Lys Ala Arg Leu Ala Ser Cys Gly Leu Ser Lys Glu
        355                 360                 365

gcg tgt gaa gat ttc cgc gca atg ata ctg atg aca ccg cac gag acg    1152
Ala Cys Glu Asp Phe Arg Ala Met Ile Leu Met Thr Pro His Glu Thr
    370                 375                 380

acg cca aag tgg gaa aat ttt acg gac ttt tta gaa atg gta aac caa    1200
Thr Pro Lys Trp Glu Asn Phe Thr Asp Phe Leu Glu Met Val Asn Gln
385                 390                 395                 400

tta aca cta tac ggg ttc tat ttt tac gag tgt ctt aac cag tac agc    1248
Leu Thr Leu Tyr Gly Phe Tyr Phe Tyr Glu Cys Leu Asn Gln Tyr Ser
                405                 410                 415

ccc act agc ata tcg ctg gcc aaa att caa aat atc tta aac cga gtc    1296
Pro Thr Ser Ile Ser Leu Ala Lys Ile Gln Asn Ile Leu Asn Arg Val
            420                 425                 430

gat gcc gaa cag agc gat cgc gcc ctg tgg cgc act cca tta atc ggg    1344
Asp Ala Glu Gln Ser Asp Arg Ala Leu Trp Arg Thr Pro Leu Ile Gly
        435                 440                 445
```

tct ttc ccg ttc ccc tgg aaa ttg aac aac gtg ctg gcg ttt ttt aag 1392
Ser Phe Pro Phe Pro Trp Lys Leu Asn Asn Val Leu Ala Phe Phe Lys
450 455 460

cct agc acc ccc gtg gca acg cta caa aaa atc tac aag gca ata ccg 1440
Pro Ser Thr Pro Val Ala Thr Leu Gln Lys Ile Tyr Lys Ala Ile Pro
465 470 475 480

tcg tac cta atg agg tct ttg ttc gaa ata gcg gcc aac aaa tcg tgg 1488
Ser Tyr Leu Met Arg Ser Leu Phe Glu Ile Ala Ala Asn Lys Ser Trp
485 490 495

ggt aac atc gcg ctc gca gag agc gcc ccg cta acc gac ata cag acc 1536
Gly Asn Ile Ala Leu Ala Glu Ser Ala Pro Leu Thr Asp Ile Gln Thr
500 505 510

gcc gaa cca gac cag ggt ccc gtg tcc gcc cag gtt atc gca aaa tat 1584
Ala Glu Pro Asp Gln Gly Pro Val Ser Ala Gln Val Ile Ala Lys Tyr
515 520 525

tgc agc cgt ctc caa ata agc gcg aca gat tat gac gca gcc atc gtc 1632
Cys Ser Arg Leu Gln Ile Ser Ala Thr Asp Tyr Asp Ala Ala Ile Val
530 535 540

tcc agc ccg gga ttc gcc gcc gag ttc att aaa aca aaa ctg tat ccg 1680
Ser Ser Pro Gly Phe Ala Ala Glu Phe Ile Lys Thr Lys Leu Tyr Pro
545 550 555 560

atc ctt agc gag gtg ctc cga aac acg tcc aaa aaa aat cga tcg cta 1728
Ile Leu Ser Glu Val Leu Arg Asn Thr Ser Lys Lys Asn Arg Ser Leu
565 570 575

ttc caa att aga tgg ctc ata gtc ttc gcg gca gag gac gcc aga gac 1776
Phe Gln Ile Arg Trp Leu Ile Val Phe Ala Ala Glu Asp Ala Arg Asp
580 585 590

cta gcc cct atc aga cgc tcg ctg gcc ctg gcg tat ttt caa ata atg 1824
Leu Ala Pro Ile Arg Arg Ser Leu Ala Leu Ala Tyr Phe Gln Ile Met
595 600 605

gac att ttg gaa gaa aag cat tct ccg gag tcg ttt tac aac ctg ctg 1872
Asp Ile Leu Glu Glu Lys His Ser Pro Glu Ser Phe Tyr Asn Leu Leu
610 615 620

gac tat ctt cag gaa aca ttt agg tgc ata cga cag gtg ata ccg gaa 1920
Asp Tyr Leu Gln Glu Thr Phe Arg Cys Ile Arg Gln Val Ile Pro Glu
625 630 635 640

gcc acg tgc cca caa gaa ttt cta caa tac ttg ttt act ttt caa aac 1968
Ala Thr Cys Pro Gln Glu Phe Leu Gln Tyr Leu Phe Thr Phe Gln Asn
645 650 655

att cca ata gca gcg agc ttc att caa aca tcc atg acc ttt gta gac 2016
Ile Pro Ile Ala Ala Ser Phe Ile Gln Thr Ser Met Thr Phe Val Asp
660 665 670

gac ctg aaa aac ggc atc ccc ggt ata ctg gac ctt gtc tcc tta ggt 2064
Asp Leu Lys Asn Gly Ile Pro Gly Ile Leu Asp Leu Val Ser Leu Gly
675 680 685

gcc gcg ttt tat aac atg aaa cta ctg tac gat tca acg cta gac acc 2112

Ala Ala Phe Tyr Asn Met Lys Leu Leu Tyr Asp Ser Thr Leu Asp Thr
690 695 700

gtc gag att cca acg gaa gaa ggg caa ccc atc gtc gtg agc atg ttc 2160
Val Glu Ile Pro Thr Glu Glu Gly Gln Pro Ile Val Val Ser Met Phe
705 710 715 720

gta ttc aaa tcc acg att cgc gtc ctc gag aag ctc tta cag gaa gcc 2208
Val Phe Lys Ser Thr Ile Arg Val Leu Glu Lys Leu Leu Gln Glu Ala
725 730 735

gtt atc gcg tta act caa acg tcc gaa ccg atg tac gca gcg cac atc 2256
Val Ile Ala Leu Thr Gln Thr Ser Glu Pro Met Tyr Ala Ala His Ile
740 745 750

cgt ctg atg caa cac ctc acg tac atg caa aag atc gcc gga cac gaa 2304
Arg Leu Met Gln His Leu Thr Tyr Met Gln Lys Ile Ala Gly His Glu
755 760 765

ata atg acc aca caa ctt cca tcc gtg ttt cac gaa ata cac gag gga 2352
Ile Met Thr Thr Gln Leu Pro Ser Val Phe His Glu Ile His Glu Gly
770 775 780

tat ttg cag tgt ttt aag cgt ttt aaa cgt ctc atg tta cac gtt acg 2400
Tyr Leu Gln Cys Phe Lys Arg Phe Lys Arg Leu Met Leu His Val Thr
785 790 795 800

gga agc tgc tgc tac tca ctg acg cga tac ttt gga ttc cta tat caa 2448
Gly Ser Cys Cys Tyr Ser Leu Thr Arg Tyr Phe Gly Phe Leu Tyr Gln
805 810 815

ccc ccc cta ata ccc gat acc atc gta caa aaa att tta aac ttt aac 2496
Pro Pro Leu Ile Pro Asp Thr Ile Val Gln Lys Ile Leu Asn Phe Asn
820 825 830

gac aaa acg gac aca acc gac gac atc tta aag agc ctg tca cag ccc 2544
Asp Lys Thr Asp Thr Thr Asp Asp Ile Leu Lys Ser Leu Ser Gln Pro
835 840 845

gtg aga caa gga cct cta tcg gct gaa aac gaa agt agc agt cga ctc 2592
Val Arg Gln Gly Pro Leu Ser Ala Glu Asn Glu Ser Ser Ser Arg Leu
850 855 860

tca aaa aac aac gtt gag ctg ctt caa aaa ctg tac gac gac ttt cgg 2640
Ser Lys Asn Asn Val Glu Leu Leu Gln Lys Leu Tyr Asp Asp Phe Arg
865 870 875 880

acc gcc tcc aca aac aat aac ccc acc tct att aaa ctt gaa tat tcg 2688
Thr Ala Ser Thr Asn Asn Asn Pro Thr Ser Ile Lys Leu Glu Tyr Ser
885 890 895

ggt aat tat aac gaa aca caa gtg tcc gta gat tgg agc aca tat aac 2736
Gly Asn Tyr Asn Glu Thr Gln Val Ser Val Asp Trp Ser Thr Tyr Asn
900 905 910

ctg gtg aca tac acc gca ccc gac gat acg tta aaa ttc acc ccg gtt 2784
Leu Val Thr Tyr Thr Ala Pro Asp Asp Thr Leu Lys Phe Thr Pro Val
915 920 925

aac acg gag gca cta gat cgc atg ttt gca gaa taa 2820
Asn Thr Glu Ala Leu Asp Arg Met Phe Ala Glu

930 935 940

<210> 141
<211> 939
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 141
Met Ala Ser Ser Ile Pro Ala Ala Arg Ala Asp Asn Gly Asp Glu Asn
1 5 10 15
Thr Gly Gly Leu Tyr Lys Leu Thr Asp Asn Leu Leu Thr Cys Thr Gly
20 25 30
Ser Leu Gln Gln Leu Lys Leu Leu Met Glu Phe Gln Leu Lys Pro Leu
35 40 45
Pro Thr Ala His Leu Leu Ser Met Pro Thr Val Thr Arg Phe Leu Asn
50 55 60
Thr Ala Phe Lys Ile Asp Asn Pro Leu Val Ser Phe Ile Gln Lys His
65 70 75 80
Pro Val Phe Phe Leu Met Arg Val Ala Arg Leu Pro Glu Pro Val Ile
85 90 95
Thr Asp His Gln Ser Ala Glu Thr Ser Thr Gly Ile Leu Ser Glu Val
100 105 110
Val Asn Val Leu Asn Thr Ala Ile Arg Lys Pro His Glu Ser Pro Ala
115 120 125
Ala Lys Asp Asn Asp Tyr Leu Asp Asn Arg Ala Ile Leu Ala Met Ile
130 135 140
Thr Glu Tyr Ile His His Val Thr Ser Arg Thr Pro Ser Gly Ile Pro
145 150 155 160
Pro Thr Pro Pro Met Gly Ile Ser His Leu Pro Cys Val Glu Gln Ile
165 170 175
Leu His Glu Thr His Arg Gln Tyr Trp Asn Leu Thr Leu Pro Glu Ser
180 185 190
Leu Phe Ile Asp Ile Gly Glu Val Ala Ser Pro Leu Gln Thr Trp Leu
195 200 205
Ile Leu Ser Tyr Cys Lys Lys Leu Gln Leu Ala Pro Pro Pro Leu Phe
210 215 220
Pro Pro Val Asp Glu Leu Ala Arg Arg Leu Val Thr Gly His His Glu
225 230 235 240
Leu Phe Val Pro Leu Ser Thr Ser Leu Glu Thr Tyr Ile Thr Met Pro
245 250 255
Val Ser Lys Gln Arg Ala Phe Glu Ile Tyr Ser Val Phe Ala Lys Ser
260 265 270
Lys Asn Ile Val Asp Gly Thr Pro Ile Leu Ala Phe Thr Asp Thr Glu
275 280 285
Leu Thr Thr Phe Thr Pro Glu Leu Leu Phe Leu Tyr Asp Phe Val Ile
290 295 300
Glu Ser Leu Cys Lys Asn Gln Ala Tyr Gly Cys Ser Arg Asn Ala Ile
305 310 315 320
Glu His Phe Ile Lys Lys Gly Ile Asp Phe Met Ala Glu Leu Gly Ala
325 330 335
Phe Ile Glu Lys Thr Cys Gly Tyr Arg Ser Thr Val Ser Leu Ser Asn
340 345 350
Val Arg Ala Val Lys Ala Arg Leu Ala Ser Cys Gly Leu Ser Lys Glu
355 360 365
Ala Cys Glu Asp Phe Arg Ala Met Ile Leu Met Thr Pro His Glu Thr
370 375 380
Thr Pro Lys Trp Glu Asn Phe Thr Asp Phe Leu Glu Met Val Asn Gln
385 390 395 400
Leu Thr Leu Tyr Gly Phe Tyr Phe Tyr Glu Cys Leu Asn Gln Tyr Ser
405 410 415

Pro Thr Ser Ile Ser Leu Ala Lys Ile Gln Asn Ile Leu Asn Arg Val
420 425 430
Asp Ala Glu Gln Ser Asp Arg Ala Leu Trp Arg Thr Pro Leu Ile Gly
435 440 445
Ser Phe Pro Phe Pro Trp Lys Leu Asn Asn Val Leu Ala Phe Phe Lys
450 455 460
Pro Ser Thr Pro Val Ala Thr Leu Gln Lys Ile Tyr Lys Ala Ile Pro
465 470 475 480
Ser Tyr Leu Met Arg Ser Leu Phe Glu Ile Ala Ala Asn Lys Ser Trp
485 490 495
Gly Asn Ile Ala Leu Ala Glu Ser Ala Pro Leu Thr Asp Ile Gln Thr
500 505 510
Ala Glu Pro Asp Gln Gly Pro Val Ser Ala Gln Val Ile Ala Lys Tyr
515 520 525
Cys Ser Arg Leu Gln Ile Ser Ala Thr Asp Tyr Asp Ala Ala Ile Val
530 535 540
Ser Ser Pro Gly Phe Ala Ala Glu Phe Ile Lys Thr Lys Leu Tyr Pro
545 550 555 560
Ile Leu Ser Glu Val Leu Arg Asn Thr Ser Lys Lys Asn Arg Ser Leu
565 570 575
Phe Gln Ile Arg Trp Leu Ile Val Phe Ala Ala Glu Asp Ala Arg Asp
580 585 590
Leu Ala Pro Ile Arg Arg Ser Leu Ala Leu Ala Tyr Phe Gln Ile Met
595 600 605
Asp Ile Leu Glu Glu Lys His Ser Pro Glu Ser Phe Tyr Asn Leu Leu
610 615 620
Asp Tyr Leu Gln Glu Thr Phe Arg Cys Ile Arg Gln Val Ile Pro Glu
625 630 635 640
Ala Thr Cys Pro Gln Glu Phe Leu Gln Tyr Leu Phe Thr Phe Gln Asn
645 650 655
Ile Pro Ile Ala Ala Ser Phe Ile Gln Thr Ser Met Thr Phe Val Asp
660 665 670
Asp Leu Lys Asn Gly Ile Pro Gly Ile Leu Asp Leu Val Ser Leu Gly
675 680 685
Ala Ala Phe Tyr Asn Met Lys Leu Leu Tyr Asp Ser Thr Leu Asp Thr
690 695 700
Val Glu Ile Pro Thr Glu Glu Gly Gln Pro Ile Val Val Ser Met Phe
705 710 715 720
Val Phe Lys Ser Thr Ile Arg Val Leu Glu Lys Leu Leu Gln Glu Ala
725 730 735
Val Ile Ala Leu Thr Gln Thr Ser Glu Pro Met Tyr Ala Ala His Ile
740 745 750
Arg Leu Met Gln His Leu Thr Tyr Met Gln Lys Ile Ala Gly His Glu
755 760 765
Ile Met Thr Thr Gln Leu Pro Ser Val Phe His Glu Ile His Glu Gly
770 775 780
Tyr Leu Gln Cys Phe Lys Arg Phe Lys Arg Leu Met Leu His Val Thr
785 790 795 800
Gly Ser Cys Cys Tyr Ser Leu Thr Arg Tyr Phe Gly Phe Leu Tyr Gln
805 810 815
Pro Pro Leu Ile Pro Asp Thr Ile Val Gln Lys Ile Leu Asn Phe Asn
820 825 830
Asp Lys Thr Asp Thr Thr Asp Asp Ile Leu Lys Ser Leu Ser Gln Pro
835 840 845
Val Arg Gln Gly Pro Leu Ser Ala Glu Asn Glu Ser Ser Ser Arg Leu
850 855 860
Ser Lys Asn Asn Val Glu Leu Leu Gln Lys Leu Tyr Asp Asp Phe Arg
865 870 875 880
Thr Ala Ser Thr Asn Asn Asn Pro Thr Ser Ile Lys Leu Glu Tyr Ser
885 890 895
Gly Asn Tyr Asn Glu Thr Gln Val Ser Val Asp Trp Ser Thr Tyr Asn

```
                900                     905                     910
Leu Val Thr Tyr Thr Ala Pro Asp Asp Thr Leu Lys Phe Thr Pro Val
            915                     920                     925
Asn Thr Glu Ala Leu Asp Arg Met Phe Ala Glu
        930                     935
```

<210> 142
<211> 7647
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(7647)

<400> 142

```
atg gaa ttg ccg cca att ttt tcg aaa ttt aaa ata gag ggc gtg gca      48
Met Glu Leu Pro Pro Ile Phe Ser Lys Phe Lys Ile Glu Gly Val Ala
1               5                   10                  15

acg acg cac cag gcc gac tgc agg ttc gga caa tac gcc ggc tcg cag      96
Thr Thr His Gln Ala Asp Cys Arg Phe Gly Gln Tyr Ala Gly Ser Gln
            20                  25                  30

tgc tta agc aac tgc gtt atc tac ctc gca caa agt tat ttc aat aga     144
Cys Leu Ser Asn Cys Val Ile Tyr Leu Ala Gln Ser Tyr Phe Asn Arg
        35                  40                  45

gag tcc ccc gtg acg gac act aac gac ctc gac gac gtt tta cgg caa     192
Glu Ser Pro Val Thr Asp Thr Asn Asp Leu Asp Asp Val Leu Arg Gln
    50                  55                  60

ggg gcg acg ttg gac ttt ata ctg aga cgg tcc ggg acg ctc ggc tat     240
Gly Ala Thr Leu Asp Phe Ile Leu Arg Arg Ser Gly Thr Leu Gly Tyr
65                  70                  75                  80

aac caa tat gcc caa ctg cac cac ata ccc agc ttt atc aag acc aac     288
Asn Gln Tyr Ala Gln Leu His His Ile Pro Ser Phe Ile Lys Thr Asn
                85                  90                  95

gag tgg acc gcg gcc atc ttt caa tct cag gag tac ttt gga cta att     336
Glu Trp Thr Ala Ala Ile Phe Gln Ser Gln Glu Tyr Phe Gly Leu Ile
            100                 105                 110

gga ctt gac gcg gcc atc cgc gaa cct ttc atc gaa tcc cta aaa tcg     384
Gly Leu Asp Ala Ala Ile Arg Glu Pro Phe Ile Glu Ser Leu Lys Ser
        115                 120                 125

atc cta acg cga aat tac gcc ggc acc gtt caa tac ttt ctg ttt ata     432
Ile Leu Thr Arg Asn Tyr Ala Gly Thr Val Gln Tyr Phe Leu Phe Ile
    130                 135                 140

tgc ggc gac aaa gcc ggg gcc gta att atc aaa aac aaa acg ttt tac     480
Cys Gly Asp Lys Ala Gly Ala Val Ile Ile Lys Asn Lys Thr Phe Tyr
145                 150                 155                 160

ctg ttc gat ccc cat tgc gta ccc cac gta cca aat agt ccg gca cac     528
Leu Phe Asp Pro His Cys Val Pro His Val Pro Asn Ser Pro Ala His
                165                 170                 175
```

gta atc agt tcg tca gac ccc acc gcc ata cta gag tac gtg tca ccg 576
Val Ile Ser Ser Ser Asp Pro Thr Ala Ile Leu Glu Tyr Val Ser Pro
180 185 190

cca gac aga gaa tac act ggc agt ttt cta tac att atg ccc agc gaa 624
Pro Asp Arg Glu Tyr Thr Gly Ser Phe Leu Tyr Ile Met Pro Ser Glu
195 200 205

tat gtc aat cca gag cac tac atc acc aac cac tat aga act ata acg 672
Tyr Val Asn Pro Glu His Tyr Ile Thr Asn His Tyr Arg Thr Ile Thr
210 215 220

ttc gcc aaa gta cac ggc ccg cac ata gat ata tcc acc ggc ata gag 720
Phe Ala Lys Val His Gly Pro His Ile Asp Ile Ser Thr Gly Ile Glu
225 230 235 240

ccg tgc acc atc gaa gac atc cca agc ccg ccg cga tcg ccg gat gtg 768
Pro Cys Thr Ile Glu Asp Ile Pro Ser Pro Pro Arg Ser Pro Asp Val
245 250 255

acg tca aaa tca tcc aac ctc gca cgc gtg ccg aga acc acc acc gac 816
Thr Ser Lys Ser Ser Asn Leu Ala Arg Val Pro Arg Thr Thr Thr Asp
260 265 270

acg tcg agc gcc aaa ccc cca ccg gcg acg cta tcc ggt tta cgg ggc 864
Thr Ser Ser Ala Lys Pro Pro Pro Ala Thr Leu Ser Gly Leu Arg Gly
275 280 285

gcg gaa cca ccg aca agc tac cca gac ccg gca acc aac gac gcg gac 912
Ala Glu Pro Pro Thr Ser Tyr Pro Asp Pro Ala Thr Asn Asp Ala Asp
290 295 300

aca aaa ctc tta act ccc gct cca gcg caa acc gcc gtc gac cac ccc 960
Thr Lys Leu Leu Thr Pro Ala Pro Ala Gln Thr Ala Val Asp His Pro
305 310 315 320

gaa ttc caa aca aca cct gga gcc acg cta ctt ctt tca gaa ctg tcg 1008
Glu Phe Gln Thr Thr Pro Gly Ala Thr Leu Leu Leu Ser Glu Leu Ser
325 330 335

gca tcc agg ggt cgg aaa cgc aag ctt tcc agc ctt cag cga tat tcg 1056
Ala Ser Arg Gly Arg Lys Arg Lys Leu Ser Ser Leu Gln Arg Tyr Ser
340 345 350

gat tcc gac gaa gcg tcg tct gac gac gaa ggg gcc cca cgt agg cgc 1104
Asp Ser Asp Glu Ala Ser Ser Asp Asp Glu Gly Ala Pro Arg Arg Arg
355 360 365

gta cac gac gat gcg ata tcc gcc gag gtg atc tgg atg gac gac gat 1152
Val His Asp Asp Ala Ile Ser Ala Glu Val Ile Trp Met Asp Asp Asp
370 375 380

ata tct cct cta tat tca ccg tcg gcg act cca tcg ttt gac gac gtg 1200
Ile Ser Pro Leu Tyr Ser Pro Ser Ala Thr Pro Ser Phe Asp Asp Val
385 390 395 400

ttc gac agc ccc cca atg agc ccg gag ttt aca tac gaa gac gcg aca 1248
Phe Asp Ser Pro Pro Met Ser Pro Glu Phe Thr Tyr Glu Asp Ala Thr
405 410 415

```
gag gac acg gat ggc gcg ttt cta gaa cag atc gct cga gac gcg gaa     1296
Glu Asp Thr Asp Gly Ala Phe Leu Glu Gln Ile Ala Arg Asp Ala Glu
            420                 425                 430

acg ccg ttc tct gcg ttc gac gac ctt ata acg gac cac gac ttt tct     1344
Thr Pro Phe Ser Ala Phe Asp Asp Leu Ile Thr Asp His Asp Phe Ser
        435                 440                 445

tcc ctt gat aaa aaa ata gaa cag tta atc aag tac gaa gcg ccc tcg     1392
Ser Leu Asp Lys Lys Ile Glu Gln Leu Ile Lys Tyr Glu Ala Pro Ser
    450                 455                 460

cag cac ctg cca aac atc tcg gac aaa caa aac ggg cga gcc gtc cga     1440
Gln His Leu Pro Asn Ile Ser Asp Lys Gln Asn Gly Arg Ala Val Arg
465                 470                 475                 480

gaa gcg gcg gcc ctc cag gcg atg gac aaa att atg att aat atc ata     1488
Glu Ala Ala Ala Leu Gln Ala Met Asp Lys Ile Met Ile Asn Ile Ile
            485                 490                 495

ctt gaa cac ggt cta att aca gac gcg cag gcc cgg gga ccg tcc gcg     1536
Leu Glu His Gly Leu Ile Thr Asp Ala Gln Ala Arg Gly Pro Ser Ala
            500                 505                 510

tgc aaa aac gtt ctt caa ttt ttc atc ctg tgg gga gaa aaa ctc aac     1584
Cys Lys Asn Val Leu Gln Phe Phe Ile Leu Trp Gly Glu Lys Leu Asn
        515                 520                 525

ata cca ata agc gac gcc aag cag gtc cta gaa ctc gat ctg caa ctg     1632
Ile Pro Ile Ser Asp Ala Lys Gln Val Leu Glu Leu Asp Leu Gln Leu
    530                 535                 540

ata ccg tta cat acg gct atc agc gaa gga aaa ttc aaa cag ggg gcg     1680
Ile Pro Leu His Thr Ala Ile Ser Glu Gly Lys Phe Lys Gln Gly Ala
545                 550                 555                 560

ttc aaa aaa cac cta aca act aaa atc aac cgg tgt ctg gcg tca atg     1728
Phe Lys Lys His Leu Thr Thr Lys Ile Asn Arg Cys Leu Ala Ser Met
                565                 570                 575

agg gcc acg cac gca gat gcg caa aaa aaa ctg gcg tcg gct ttc aac     1776
Arg Ala Thr His Ala Asp Ala Gln Lys Lys Leu Ala Ser Ala Phe Asn
            580                 585                 590

gtc gaa ggc tca cag att tcg tcc agc gaa gcg aaa ata tca gtc cgg     1824
Val Glu Gly Ser Gln Ile Ser Ser Ser Glu Ala Lys Ile Ser Val Arg
        595                 600                 605

gcg ctg aag gaa cag atc gcc aac cac ctg agt cca ggc ttt tta gcg     1872
Ala Leu Lys Glu Gln Ile Ala Asn His Leu Ser Pro Gly Phe Leu Ala
    610                 615                 620

gtc tac tcc gcg gac gag gta aaa cat cta cgg gat aaa att cag gac     1920
Val Tyr Ser Ala Asp Glu Val Lys His Leu Arg Asp Lys Ile Gln Asp
625                 630                 635                 640

cta aaa aca ggc atc gag cag cgc aac aaa gaa atc caa cag gaa gaa     1968
Leu Lys Thr Gly Ile Glu Gln Arg Asn Lys Glu Ile Gln Gln Glu Glu
                645                 650                 655

ctg ttt ttt gat gcc atg ctc aca gcc ctg gac acg ttc caa ccc cct     2016
```

Leu Phe Phe Asp Ala Met Leu Thr Ala Leu Asp Thr Phe Gln Pro Pro
660 665 670

ccg aaa acg gca ttt cca atg gag atc ttt ccg cac cgt aaa acc gaa 2064
Pro Lys Thr Ala Phe Pro Met Glu Ile Phe Pro His Arg Lys Thr Glu
675 680 685

gtt atg ctc gac cac ctg gcg tcc ata acc acc agg tta acc gag gac 2112
Val Met Leu Asp His Leu Ala Ser Ile Thr Thr Arg Leu Thr Glu Asp
690 695 700

gcc acc gaa gcc ctc aac aat tac ctg gag acc ccg ccc gac caa gga 2160
Ala Thr Glu Ala Leu Asn Asn Tyr Leu Glu Thr Pro Pro Asp Gln Gly
705 710 715 720

acg cac att acc aac att cca aac ttt tca tcc atc gtg gca aat atc 2208
Thr His Ile Thr Asn Ile Pro Asn Phe Ser Ser Ile Val Ala Asn Ile
725 730 735

ata tcc acg tta aaa atc cta acg tac gca gaa aac gac atg caa tta 2256
Ile Ser Thr Leu Lys Ile Leu Thr Tyr Ala Glu Asn Asp Met Gln Leu
740 745 750

aac gta acg ccc atg gca acg tac agg cgt cag ctg ttg tac ctc gga 2304
Asn Val Thr Pro Met Ala Thr Tyr Arg Arg Gln Leu Leu Tyr Leu Gly
755 760 765

ggc gag tta gca acc atc ttt aat tta gag tgg cca tac gaa acc gtg 2352
Gly Glu Leu Ala Thr Ile Phe Asn Leu Glu Trp Pro Tyr Glu Thr Val
770 775 780

cca ccg gtt caa gaa ctg ccc ctc gtg gcg cgg gcg aaa gca aaa atg 2400
Pro Pro Val Gln Glu Leu Pro Leu Val Ala Arg Ala Lys Ala Lys Met
785 790 795 800

gaa tcg gta aca aaa atg gaa aag aac caa cag gct ctc gac caa ata 2448
Glu Ser Val Thr Lys Met Glu Lys Asn Gln Gln Ala Leu Asp Gln Ile
805 810 815

ctg gga gac gcc gaa acg tta ctt gac aca ata acc gca aca tcc gga 2496
Leu Gly Asp Ala Glu Thr Leu Leu Asp Thr Ile Thr Ala Thr Ser Gly
820 825 830

gat gag aac ccg gtc cgc gcc atg tcc ata ccg ata ctg gag acc tac 2544
Asp Glu Asn Pro Val Arg Ala Met Ser Ile Pro Ile Leu Glu Thr Tyr
835 840 845

att aca aac gca ggc gcc ctg ata ggc agt tct cga aac cag cgg ttc 2592
Ile Thr Asn Ala Gly Ala Leu Ile Gly Ser Ser Arg Asn Gln Arg Phe
850 855 860

gaa aaa ctc aag gcc gcc atc cac gac ctg gca tca tcg gag tcg ttc 2640
Glu Lys Leu Lys Ala Ala Ile His Asp Leu Ala Ser Ser Glu Ser Phe
865 870 875 880

ata ata atg ctg cta aac aac acg cgg ctc gat aac ata tca gac aat 2688
Ile Ile Met Leu Leu Asn Asn Thr Arg Leu Asp Asn Ile Ser Asp Asn
885 890 895

ctg gcc aag atc gac ggc atc ctg acc aac aac aca cgt ttt ctt tca 2736
Leu Ala Lys Ile Asp Gly Ile Leu Thr Asn Asn Thr Arg Phe Leu Ser

900 905 910

aac gcc act gtt agc aaa acg ctc caa acg ctg gga ggc agc cta ata 2784
Asn Ala Thr Val Ser Lys Thr Leu Gln Thr Leu Gly Gly Ser Leu Ile
915 920 925

cgc gaa tgc gta gaa gcg cta aat aaa agg agc ccc tct tcc ctc aac 2832
Arg Glu Cys Val Glu Ala Leu Asn Lys Arg Ser Pro Ser Ser Leu Asn
930 935 940

aac gcg cgt ctc ctc gcg gtt caa acc ata ctg ggg cac gcg tcc gtt 2880
Asn Ala Arg Leu Leu Ala Val Gln Thr Ile Leu Gly His Ala Ser Val
945 950 955 960

cca gat cac gag acg ctg acg cga atc gtt tcc ggc gtc gcc agc gca 2928
Pro Asp His Glu Thr Leu Thr Arg Ile Val Ser Gly Val Ala Ser Ala
965 970 975

caa aag gaa tcc gct ggc gat gat cca gat agg tgg acg cga gta acc 2976
Gln Lys Glu Ser Ala Gly Asp Asp Pro Asp Arg Trp Thr Arg Val Thr
980 985 990

ggt cac cta aac gag ctg aag ctc gta act acc caa tcg cgt gtc gac 3024
Gly His Leu Asn Glu Leu Lys Leu Val Thr Thr Gln Ser Arg Val Asp
995 1000 1005

aaa gcc acc agg cgc aaa ctg tta atg ata ata acc cgt gac ctc aag 3072
Lys Ala Thr Arg Arg Lys Leu Leu Met Ile Ile Thr Arg Asp Leu Lys
1010 1015 1020

gag gcg gag gtg tct cag gaa acg gtc ctg gaa aca cgg tgg caa gaa 3120
Glu Ala Glu Val Ser Gln Glu Thr Val Leu Glu Thr Arg Trp Gln Glu
1025 1030 1035 1040

aac gtg cta aag ttt caa ccg tcg acg tcc aaa gaa atc gaa gac ttt 3168
Asn Val Leu Lys Phe Gln Pro Ser Thr Ser Lys Glu Ile Glu Asp Phe
1045 1050 1055

tta cag tcg gca ccg tca gca aag gcc cga aaa ttc gca gaa aaa cac 3216
Leu Gln Ser Ala Pro Ser Ala Lys Ala Arg Lys Phe Ala Glu Lys His
1060 1065 1070

cta cgg acg ctg atc acc caa ttc aac ggc cac gag cga ccg ccg tcc 3264
Leu Arg Thr Leu Ile Thr Gln Phe Asn Gly His Glu Arg Pro Pro Ser
1075 1080 1085

gag gcc acc gcc gtt ccc atg gac tac acg ccg acg ccc ata ccc acg 3312
Glu Ala Thr Ala Val Pro Met Asp Tyr Thr Pro Thr Pro Ile Pro Thr
1090 1095 1100

cca cag gcc gtt tct acg gct acc gcg gaa aag gga aag gcc gca tgg 3360
Pro Gln Ala Val Ser Thr Ala Thr Ala Glu Lys Gly Lys Ala Ala Trp
1105 1110 1115 1120

aat aaa att caa cag gcc ttt cag gat ttc aac ttt cac ctc atc gac 3408
Asn Lys Ile Gln Gln Ala Phe Gln Asp Phe Asn Phe His Leu Ile Asp
1125 1130 1135

gct tcg gat tgg caa gag atg gca tca gaa tac tcc aga cac ggc tcg 3456
Ala Ser Asp Trp Gln Glu Met Ala Ser Glu Tyr Ser Arg His Gly Ser
1140 1145 1150

tcc ctt cct ggt acg gtt gga cca aag ctg gtg cgc ttc atg gag agc 3504
Ser Leu Pro Gly Thr Val Gly Pro Lys Leu Val Arg Phe Met Glu Ser
1155 1160 1165

atc tca aac acc ctg gac gac atc ctc acg cag aag ctg gca tct ctg 3552
Ile Ser Asn Thr Leu Asp Asp Ile Leu Thr Gln Lys Leu Ala Ser Leu
1170 1175 1180

ctt cca aac ggg ccc gcg ttc aga ccc cca gcg ttt gac tgg atc gcg 3600
Leu Pro Asn Gly Pro Ala Phe Arg Pro Pro Ala Phe Asp Trp Ile Ala
1185 1190 1195 1200

ccc tat caa aca cgc gta aac gcg ttt cta aaa acc ata ggc ctg ccc 3648
Pro Tyr Gln Thr Arg Val Asn Ala Phe Leu Lys Thr Ile Gly Leu Pro
1205 1210 1215

atg gtg cgc aac ctg gcg gac aag atc cat cac caa tgc caa act gtc 3696
Met Val Arg Asn Leu Ala Asp Lys Ile His His Gln Cys Gln Thr Val
1220 1225 1230

agt cac gcg gtg caa tcc gca gac ctt caa cag gcc acg gtg gga aca 3744
Ser His Ala Val Gln Ser Ala Asp Leu Gln Gln Ala Thr Val Gly Thr
1235 1240 1245

agt tta gaa cga ccc gcg gcc gaa tac tgt cga ata ctc tct gac atg 3792
Ser Leu Glu Arg Pro Ala Ala Glu Tyr Cys Arg Ile Leu Ser Asp Met
1250 1255 1260

caa gtc gcg ttc aac gac cac gga atc gcc gta aga tcg gag gcc gcg 3840
Gln Val Ala Phe Asn Asp His Gly Ile Ala Val Arg Ser Glu Ala Ala
1265 1270 1275 1280

gcg tac acg gac gca atc aac tcg ccg gcc aac gtc gtg act ccc ccg 3888
Ala Tyr Thr Asp Ala Ile Asn Ser Pro Ala Asn Val Val Thr Pro Pro
1285 1290 1295

aaa ccc aac cta gaa gcc ccc aag aag cta ata acg gca act gac gcc 3936
Lys Pro Asn Leu Glu Ala Pro Lys Lys Leu Ile Thr Ala Thr Asp Ala
1300 1305 1310

cta acc gtc gag gac ttt cca gat ttc cta aaa acg tca atc ctt caa 3984
Leu Thr Val Glu Asp Phe Pro Asp Phe Leu Lys Thr Ser Ile Leu Gln
1315 1320 1325

cag gag cag cga ctc att gcg ctc cag aga gcg gaa ttt cag caa cta 4032
Gln Glu Gln Arg Leu Ile Ala Leu Gln Arg Ala Glu Phe Gln Gln Leu
1330 1335 1340

gag gcc agc atc tcg gcg gcc gaa cgg ctc cgc caa tcc acc cgt gac 4080
Glu Ala Ser Ile Ser Ala Ala Glu Arg Leu Arg Gln Ser Thr Arg Asp
1345 1350 1355 1360

gag atc gca ggc aag atg gca acc gct atc acg caa ctc tta ccc cgc 4128
Glu Ile Ala Gly Lys Met Ala Thr Ala Ile Thr Gln Leu Leu Pro Arg
1365 1370 1375

gcc ccc gtc gca ata tca tcg aga ccg ttg aac tta tca aaa cct ata 4176
Ala Pro Val Ala Ile Ser Ser Arg Pro Leu Asn Leu Ser Lys Pro Ile
1380 1385 1390

gac ttt ttg agt tca acg gta tac gac aaa atc ctg gac aag gag cct 4224
Asp Phe Leu Ser Ser Thr Val Tyr Asp Lys Ile Leu Asp Lys Glu Pro
1395 1400 1405

tac gag aca gcc ata gcg gga ttc gcg tgg ctg gaa atc gcg aca aaa 4272
Tyr Glu Thr Ala Ile Ala Gly Phe Ala Trp Leu Glu Ile Ala Thr Lys
1410 1415 1420

tcc gta atg gtc tac agt caa caa aac gaa acg caa cag tta aac gta 4320
Ser Val Met Val Tyr Ser Gln Gln Asn Glu Thr Gln Gln Leu Asn Val
1425 1430 1435 1440

ctg ctg agc gag gta gaa aaa cag agc acc gtc gcg cag cgt cta cac 4368
Leu Leu Ser Glu Val Glu Lys Gln Ser Thr Val Ala Gln Arg Leu His
1445 1450 1455

gat ttg gaa ctg tcg gcg aaa aac acg gac gac gta aag gtg ctg aag 4416
Asp Leu Glu Leu Ser Ala Lys Asn Thr Asp Asp Val Lys Val Leu Lys
1460 1465 1470

cag gcg cta gac gaa ctc gcg ccc ctc agg gta aag ggc gga aaa acc 4464
Gln Ala Leu Asp Glu Leu Ala Pro Leu Arg Val Lys Gly Gly Lys Thr
1475 1480 1485

acc gta gac gcg tgg aaa caa aaa ctg gaa agc ata gaa tcc ctg ctt 4512
Thr Val Asp Ala Trp Lys Gln Lys Leu Glu Ser Ile Glu Ser Leu Leu
1490 1495 1500

cgc gcc acg agg acg gca ggc gaa ata tca tcg gag ctt gaa cgc atc 4560
Arg Ala Thr Arg Thr Ala Gly Glu Ile Ser Ser Glu Leu Glu Arg Ile
1505 1510 1515 1520

ggc aca cag gcg gtt ggc acc atc acc gtc cgc gat tta gga acg ctc 4608
Gly Thr Gln Ala Val Gly Thr Ile Thr Val Arg Asp Leu Gly Thr Leu
1525 1530 1535

tcc gat caa tgc cgg gaa gcc gca aat ttc ctc aga cag gcc agt cta 4656
Ser Asp Gln Cys Arg Glu Ala Ala Asn Phe Leu Arg Gln Ala Ser Leu
1540 1545 1550

ccc gaa ggc ttc tcg gac ata ggc aca aaa ctc agc gag ctt cag gcg 4704
Pro Glu Gly Phe Ser Asp Ile Gly Thr Lys Leu Ser Glu Leu Gln Ala
1555 1560 1565

tac att aag tac aaa aaa cag ttt ctg gag cat ttt gaa aca acc cag 4752
Tyr Ile Lys Tyr Lys Lys Gln Phe Leu Glu His Phe Glu Thr Thr Gln
1570 1575 1580

cct aac gtc ttt caa cgc ttc ccg cta tcc caa aac ata acc gaa aac 4800
Pro Asn Val Phe Gln Arg Phe Pro Leu Ser Gln Asn Ile Thr Glu Asn
1585 1590 1595 1600

gtt ccg gcg cgc ccg gcg atg gac tcg gtg gcc aga ctg acc aat cac 4848
Val Pro Ala Arg Pro Ala Met Asp Ser Val Ala Arg Leu Thr Asn His
1605 1610 1615

ctt cac gtg cgc ggc agc gcg ccc cac ttt aca acg tgg ata gaa acg 4896
Leu His Val Arg Gly Ser Ala Pro His Phe Thr Thr Trp Ile Glu Thr
1620 1625 1630

cta ccg acc gtc gat ccg gaa aaa cca act cac gtc ccg gcg cac gga 4944

```
Leu Pro Thr Val Asp Pro Glu Lys Pro Thr His Val Pro Ala His Gly
        1635                1640                1645

gga gcc cct ctg cac cgc cag atc acg tac tca aac gtc cta gag gcg    4992
Gly Ala Pro Leu His Arg Gln Ile Thr Tyr Ser Asn Val Leu Glu Ala
    1650                1655                1660

ttg ttt tca tta tgt tcc acc acg cta acc ccg gtt ccg acg gcc ccc    5040
Leu Phe Ser Leu Cys Ser Thr Thr Leu Thr Pro Val Pro Thr Ala Pro
1665                1670                1675                1680

ggt ctg gaa atc gca acc agg gca cgc cgc ggg gca gag gcc gca acg    5088
Gly Leu Glu Ile Ala Thr Arg Ala Arg Arg Gly Ala Glu Ala Ala Thr
                1685                1690                1695

tgg atg gac aga cag tgg ccc gac atc gct cag acg ctc caa gac gtt    5136
Trp Met Asp Arg Gln Trp Pro Asp Ile Ala Gln Thr Leu Gln Asp Val
            1700                1705                1710

ctc gac acg tac gaa cac acc acc gcc cac gca aac cgg gac gcg gca    5184
Leu Asp Thr Tyr Glu His Thr Thr Ala His Ala Asn Arg Asp Ala Ala
        1715                1720                1725

ttt aac aca ttc ttg gcg atg tgc gtt ttt acg caa atc atc agg ggc    5232
Phe Asn Thr Phe Leu Ala Met Cys Val Phe Thr Gln Ile Ile Arg Gly
    1730                1735                1740

gct agc aga gcc gtg acg ctc ccg aag tta ccg agc acc gcc gtc gat    5280
Ala Ser Arg Ala Val Thr Leu Pro Lys Leu Pro Ser Thr Ala Val Asp
1745                1750                1755                1760

ttt cca gaa gag atc gtt cta aca ccc agg gaa tgc aca aca ctg gtc    5328
Phe Pro Glu Glu Ile Val Leu Thr Pro Arg Glu Cys Thr Thr Leu Val
                1765                1770                1775

acc gcc atg tgg ccc acc ctg gcg gcc gca atc tta cga tta aaa tcc    5376
Thr Ala Met Trp Pro Thr Leu Ala Ala Ala Ile Leu Arg Leu Lys Ser
            1780                1785                1790

tac tcg gaa gcc cta gga cta atg agt cgt ttc ctc ccg cta atg ttc    5424
Tyr Ser Glu Ala Leu Gly Leu Met Ser Arg Phe Leu Pro Leu Met Phe
        1795                1800                1805

cag gcg ctg ccg cac cta acg cta gag gcc cag gtc aaa aac ggc cca    5472
Gln Ala Leu Pro His Leu Thr Leu Glu Ala Gln Val Lys Asn Gly Pro
    1810                1815                1820

cat aac acc ccg cct cag ttg aga tgc ttt gcc aaa aca gag gca att    5520
His Asn Thr Pro Pro Gln Leu Arg Cys Phe Ala Lys Thr Glu Ala Ile
1825                1830                1835                1840

ccg tat ttc ccg gcg caa tgg cag tca gcg aac cta gag cag agc ctg    5568
Pro Tyr Phe Pro Ala Gln Trp Gln Ser Ala Asn Leu Glu Gln Ser Leu
                1845                1850                1855

tgg gga cag acg gac ttt ttg caa atc tgc gat aac aat caa cgc aag    5616
Trp Gly Gln Thr Asp Phe Leu Gln Ile Cys Asp Asn Asn Gln Arg Lys
            1860                1865                1870

gcc agg gtg gcg gcc gtc acc tgg gcg ctc acg acg ata gac ggc gtg    5664
Ala Arg Val Ala Ala Val Thr Trp Ala Leu Thr Thr Ile Asp Gly Val
```

1875 1880 1885

gtt ttg gac caa ctg tgg tcc aca ttt aaa ccc atg aca gcc gcg tca 5712
Val Leu Asp Gln Leu Trp Ser Thr Phe Lys Pro Met Thr Ala Ala Ser
1890 1895 1900

gac gac acg tac gtc gac cta gtc gag acc cta cac ctg acc acc ttt 5760
Asp Asp Thr Tyr Val Asp Leu Val Glu Thr Leu His Leu Thr Thr Phe
1905 1910 1915 1920

ggc ccg cgc ggt cca acg ccg agg cga gaa acg acc acc gag cac ccg 5808
Gly Pro Arg Gly Pro Thr Pro Arg Arg Glu Thr Thr Thr Glu His Pro
1925 1930 1935

ccg tac gag tac gga cag ccc acg ggc tac tgc atc tcg ggt caa tcg 5856
Pro Tyr Glu Tyr Gly Gln Pro Thr Gly Tyr Cys Ile Ser Gly Gln Ser
1940 1945 1950

acg acg ccg gtc cag gct tca aac aca ccg gta tcc gct ttc gag gcg 5904
Thr Thr Pro Val Gln Ala Ser Asn Thr Pro Val Ser Ala Phe Glu Ala
1955 1960 1965

gtg ctc gga gca atg gtg ttt cac gta ccg atc aga ata ttt ttg gcg 5952
Val Leu Gly Ala Met Val Phe His Val Pro Ile Arg Ile Phe Leu Ala
1970 1975 1980

gcc acg ccc aag cgc ctt ggc cag gcg cgc ggc ggc atg ggg ctc ctc 6000
Ala Thr Pro Lys Arg Leu Gly Gln Ala Arg Gly Gly Met Gly Leu Leu
1985 1990 1995 2000

aca ccc atc ctg gaa tgc gtc ccc gac gtc gag ccc ttc aaa agc ctg 6048
Thr Pro Ile Leu Glu Cys Val Pro Asp Val Glu Pro Phe Lys Ser Leu
2005 2010 2015

tat aac gca ccc cgc aaa ccc gtg ccc att gaa acg cta ccc gca tcc 6096
Tyr Asn Ala Pro Arg Lys Pro Val Pro Ile Glu Thr Leu Pro Ala Ser
2020 2025 2030

ctc cac ccg cac gac gag cga cag gtc ttt ctg aga cag gca cag tgg 6144
Leu His Pro His Asp Glu Arg Gln Val Phe Leu Arg Gln Ala Gln Trp
2035 2040 2045

cta tcc tac cga ttc aca cca cac gaa gcc gcc cgg tcg tcg act ccg 6192
Leu Ser Tyr Arg Phe Thr Pro His Glu Ala Ala Arg Ser Ser Thr Pro
2050 2055 2060

ccg ctt ctg gtg gtc ata gac cct gaa aac ctc gta acg gca acg tac 6240
Pro Leu Leu Val Val Ile Asp Pro Glu Asn Leu Val Thr Ala Thr Tyr
2065 2070 2075 2080

tcc agt ggc ggg cct gca aat ttc gag agc agg ccg ttt tac gtg atg 6288
Ser Ser Gly Gly Pro Ala Asn Phe Glu Ser Arg Pro Phe Tyr Val Met
2085 2090 2095

ccc gga cca tac ccc cca gac tgg cca aaa acg ctg tcg gta aca tca 6336
Pro Gly Pro Tyr Pro Pro Asp Trp Pro Lys Thr Leu Ser Val Thr Ser
2100 2105 2110

aac acg tcc gtg acg cac ctc agc cac gac gag ata tgt aac ctc ttt 6384
Asn Thr Ser Val Thr His Leu Ser His Asp Glu Ile Cys Asn Leu Phe
2115 2120 2125

```
act acg cta tcc cga gaa cac ggg acc gtg caa ggc agg gat atc ttc    6432
Thr Thr Leu Ser Arg Glu His Gly Thr Val Gln Gly Arg Asp Ile Phe
    2130                2135                2140

gca gcg gct ccg aca aac gtc aca ccg gaa caa acc gcc aat cct ccg    6480
Ala Ala Ala Pro Thr Asn Val Thr Pro Glu Gln Thr Ala Asn Pro Pro
2145                2150                2155                2160

gca tgg gaa acg gat aac cga tta ata acg caa aca gaa acc gcc aaa    6528
Ala Trp Glu Thr Asp Asn Arg Leu Ile Thr Gln Thr Glu Thr Ala Lys
            2165                2170                2175

aaa cct cat ata att cct gcg tct cct aaa gcg cgg aca gat cca ccg    6576
Lys Pro His Ile Ile Pro Ala Ser Pro Lys Ala Arg Thr Asp Pro Pro
        2180                2185                2190

gtg gaa acc acg acc cac cat tca caa ggg caa gcg tcg caa cac gca    6624
Val Glu Thr Thr Thr His His Ser Gln Gly Gln Ala Ser Gln His Ala
    2195                2200                2205

aac agc aac gta aac cag ccc ggt caa att act tca cac gcg tca cgt    6672
Asn Ser Asn Val Asn Gln Pro Gly Gln Ile Thr Ser His Ala Ser Arg
    2210                2215                2220

aac aca ccg tca acc gca cct cag gcc tca tct tca ccg gaa aaa ttc    6720
Asn Thr Pro Ser Thr Ala Pro Gln Ala Ser Ser Ser Pro Glu Lys Phe
2225                2230                2235                2240

aac acg caa acg gtg cct cga cta ata tct caa acg tcg gaa acg gcc    6768
Asn Thr Gln Thr Val Pro Arg Leu Ile Ser Gln Thr Ser Glu Thr Ala
            2245                2250                2255

cat ata aac cag cca gcc tcc ggc cag gtc acc gaa cca aag gga atc    6816
His Ile Asn Gln Pro Ala Ser Gly Gln Val Thr Glu Pro Lys Gly Ile
        2260                2265                2270

ttt ggg acg tat aaa ccc cga gtg ctc acc gaa ccc gcc aaa ccc gca    6864
Phe Gly Thr Tyr Lys Pro Arg Val Leu Thr Glu Pro Ala Lys Pro Ala
    2275                2280                2285

aac gcc ggc gta gcc tct cgc caa cca gag gca acc acc acg gtc ccc    6912
Asn Ala Gly Val Ala Ser Arg Gln Pro Glu Ala Thr Thr Thr Val Pro
    2290                2295                2300

aag tta ccg att aat cca ccc acc gct agg gtc ttt ata ggg acc gcg    6960
Lys Leu Pro Ile Asn Pro Pro Thr Ala Arg Val Phe Ile Gly Thr Ala
2305                2310                2315                2320

tcc aaa ctc tcg cca gcc gtc gaa gag agc cac ggc gcc aca ccc gac    7008
Ser Lys Leu Ser Pro Ala Val Glu Glu Ser His Gly Ala Thr Pro Asp
            2325                2330                2335

gca cat cag tcg aag ata gat cgg gaa aaa tac gcc gag agt cgg cct    7056
Ala His Gln Ser Lys Ile Asp Arg Glu Lys Tyr Ala Glu Ser Arg Pro
        2340                2345                2350

cgc cgc acc cca cac ctc gaa gag ggg cca cgg gag cct cac gtc aac    7104
Arg Arg Thr Pro His Leu Glu Glu Gly Pro Arg Glu Pro His Val Asn
    2355                2360                2365
```

act cca acc agc gca cac ata aac gtc ccc tct agc caa ggt caa aaa 7152
Thr Pro Thr Ser Ala His Ile Asn Val Pro Ser Ser Gln Gly Gln Lys
2370 2375 2380

aca gta cac ggg cgc gaa aat ccc ggc ctt caa aca gca act ccc agc 7200
Thr Val His Gly Arg Glu Asn Pro Gly Leu Gln Thr Ala Thr Pro Ser
2385 2390 2395 2400

gcc ccc caa cca acc gca tca aac ccg cgc att caa tac acg ctc ccc 7248
Ala Pro Gln Pro Thr Ala Ser Asn Pro Arg Ile Gln Tyr Thr Leu Pro
2405 2410 2415

aga acg gac ggc cgg ttg ctt cac gac gaa tcg gag gtg gaa tcg acc 7296
Arg Thr Asp Gly Arg Leu Leu His Asp Glu Ser Glu Val Glu Ser Thr
2420 2425 2430

cca acc gag gag gta aaa cga tcg cca aaa aca caa gat gtg tct cac 7344
Pro Thr Glu Glu Val Lys Arg Ser Pro Lys Thr Gln Asp Val Ser His
2435 2440 2445

ggg ccc gaa ccg gac gac tcc agg tgg acc gcc ccg ctc ggt cca acc 7392
Gly Pro Glu Pro Asp Asp Ser Arg Trp Thr Ala Pro Leu Gly Pro Thr
2450 2455 2460

ata gag att cat cga ctg gaa cac ccc caa att ctc aaa aat ata aca 7440
Ile Glu Ile His Arg Leu Glu His Pro Gln Ile Leu Lys Asn Ile Thr
2465 2470 2475 2480

tca ctc acc gtc ccc act ccc aga gtc acc cca atc cct ccc act aac 7488
Ser Leu Thr Val Pro Thr Pro Arg Val Thr Pro Ile Pro Pro Thr Asn
2485 2490 2495

atc tgg ata ccc cta tcc cac gtc aac atc caa cac gaa gaa atc aca 7536
Ile Trp Ile Pro Leu Ser His Val Asn Ile Gln His Glu Glu Ile Thr
2500 2505 2510

cga gcc aag aat gtg tta atg cga ttt att caa aac gta cga aga aaa 7584
Arg Ala Lys Asn Val Leu Met Arg Phe Ile Gln Asn Val Arg Arg Lys
2515 2520 2525

ctt caa gcg tcg tct gac gct cta tcc gag gct att gcc aga ata aag 7632
Leu Gln Ala Ser Ser Asp Ala Leu Ser Glu Ala Ile Ala Arg Ile Lys
2530 2535 2540

ttt tta tat ctg taa 7647
Phe Leu Tyr Leu
2545

<210> 143
<211> 2548
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 143
Met Glu Leu Pro Pro Ile Phe Ser Lys Phe Lys Ile Glu Gly Val Ala
1 5 10 15
Thr Thr His Gln Ala Asp Cys Arg Phe Gly Gln Tyr Ala Gly Ser Gln
20 25 30
Cys Leu Ser Asn Cys Val Ile Tyr Leu Ala Gln Ser Tyr Phe Asn Arg
35 40 45

Glu Ser Pro Val Thr Asp Thr Asn Asp Leu Asp Asp Val Leu Arg Gln
50 55 60
Gly Ala Thr Leu Asp Phe Ile Leu Arg Arg Ser Gly Thr Leu Gly Tyr
65 70 75 80
Asn Gln Tyr Ala Gln Leu His His Ile Pro Ser Phe Ile Lys Thr Asn
85 90 95
Glu Trp Thr Ala Ala Ile Phe Gln Ser Gln Glu Tyr Phe Gly Leu Ile
100 105 110
Gly Leu Asp Ala Ala Ile Arg Glu Pro Phe Ile Glu Ser Leu Lys Ser
115 120 125
Ile Leu Thr Arg Asn Tyr Ala Gly Thr Val Gln Tyr Phe Leu Phe Ile
130 135 140
Cys Gly Asp Lys Ala Gly Ala Val Ile Ile Lys Asn Lys Thr Phe Tyr
145 150 155 160
Leu Phe Asp Pro His Cys Val Pro His Val Pro Asn Ser Pro Ala His
165 170 175
Val Ile Ser Ser Ser Asp Pro Thr Ala Ile Leu Glu Tyr Val Ser Pro
180 185 190
Pro Asp Arg Glu Tyr Thr Gly Ser Phe Leu Tyr Ile Met Pro Ser Glu
195 200 205
Tyr Val Asn Pro Glu His Tyr Ile Thr Asn His Tyr Arg Thr Ile Thr
210 215 220
Phe Ala Lys Val His Gly Pro His Ile Asp Ile Ser Thr Gly Ile Glu
225 230 235 240
Pro Cys Thr Ile Glu Asp Ile Pro Ser Pro Pro Arg Ser Pro Asp Val
245 250 255
Thr Ser Lys Ser Ser Asn Leu Ala Arg Val Pro Arg Thr Thr Thr Asp
260 265 270
Thr Ser Ser Ala Lys Pro Pro Pro Ala Thr Leu Ser Gly Leu Arg Gly
275 280 285
Ala Glu Pro Pro Thr Ser Tyr Pro Asp Pro Ala Thr Asn Asp Ala Asp
290 295 300
Thr Lys Leu Leu Thr Pro Ala Pro Ala Gln Thr Ala Val Asp His Pro
305 310 315 320
Glu Phe Gln Thr Thr Pro Gly Ala Thr Leu Leu Leu Ser Glu Leu Ser
325 330 335
Ala Ser Arg Gly Arg Lys Arg Lys Leu Ser Ser Leu Gln Arg Tyr Ser
340 345 350
Asp Ser Asp Glu Ala Ser Ser Asp Asp Glu Gly Ala Pro Arg Arg Arg
355 360 365
Val His Asp Asp Ala Ile Ser Ala Glu Val Ile Trp Met Asp Asp Asp
370 375 380
Ile Ser Pro Leu Tyr Ser Pro Ser Ala Thr Pro Ser Phe Asp Asp Val
385 390 395 400
Phe Asp Ser Pro Pro Met Ser Pro Glu Phe Thr Tyr Glu Asp Ala Thr
405 410 415
Glu Asp Thr Asp Gly Ala Phe Leu Glu Gln Ile Ala Arg Asp Ala Glu
420 425 430
Thr Pro Phe Ser Ala Phe Asp Asp Leu Ile Thr Asp His Asp Phe Ser
435 440 445
Ser Leu Asp Lys Lys Ile Glu Gln Leu Ile Lys Tyr Glu Ala Pro Ser
450 455 460
Gln His Leu Pro Asn Ile Ser Asp Lys Gln Asn Gly Arg Ala Val Arg
465 470 475 480
Glu Ala Ala Ala Leu Gln Ala Met Asp Lys Ile Met Ile Asn Ile Ile
485 490 495
Leu Glu His Gly Leu Ile Thr Asp Ala Gln Ala Arg Gly Pro Ser Ala
500 505 510
Cys Lys Asn Val Leu Gln Phe Phe Ile Leu Trp Gly Glu Lys Leu Asn
515 520 525
Ile Pro Ile Ser Asp Ala Lys Gln Val Leu Glu Leu Asp Leu Gln Leu

530 535 540
Ile Pro Leu His Thr Ala Ile Ser Glu Gly Lys Phe Lys Gln Gly Ala
545 550 555 560
Phe Lys Lys His Leu Thr Thr Lys Ile Asn Arg Cys Leu Ala Ser Met
565 570 575
Arg Ala Thr His Ala Asp Ala Gln Lys Lys Leu Ala Ser Ala Phe Asn
580 585 590
Val Glu Gly Ser Gln Ile Ser Ser Ser Glu Ala Lys Ile Ser Val Arg
595 600 605
Ala Leu Lys Glu Gln Ile Ala Asn His Leu Ser Pro Gly Phe Leu Ala
610 615 620
Val Tyr Ser Ala Asp Glu Val Lys His Leu Arg Asp Lys Ile Gln Asp
625 630 635 640
Leu Lys Thr Gly Ile Glu Gln Arg Asn Lys Glu Ile Gln Gln Glu Glu
645 650 655
Leu Phe Phe Asp Ala Met Leu Thr Ala Leu Asp Thr Phe Gln Pro Pro
660 665 670
Pro Lys Thr Ala Phe Pro Met Glu Ile Phe Pro His Arg Lys Thr Glu
675 680 685
Val Met Leu Asp His Leu Ala Ser Ile Thr Thr Arg Leu Thr Glu Asp
690 695 700
Ala Thr Glu Ala Leu Asn Asn Tyr Leu Glu Thr Pro Pro Asp Gln Gly
705 710 715 720
Thr His Ile Thr Asn Ile Pro Asn Phe Ser Ser Ile Val Ala Asn Ile
725 730 735
Ile Ser Thr Leu Lys Ile Leu Thr Tyr Ala Glu Asn Asp Met Gln Leu
740 745 750
Asn Val Thr Pro Met Ala Thr Tyr Arg Arg Gln Leu Leu Tyr Leu Gly
755 760 765
Gly Glu Leu Ala Thr Ile Phe Asn Leu Glu Trp Pro Tyr Glu Thr Val
770 775 780
Pro Pro Val Gln Glu Leu Pro Leu Val Ala Arg Ala Lys Ala Lys Met
785 790 795 800
Glu Ser Val Thr Lys Met Glu Lys Asn Gln Gln Ala Leu Asp Gln Ile
805 810 815
Leu Gly Asp Ala Glu Thr Leu Leu Asp Thr Ile Thr Ala Thr Ser Gly
820 825 830
Asp Glu Asn Pro Val Arg Ala Met Ser Ile Pro Ile Leu Glu Thr Tyr
835 840 845
Ile Thr Asn Ala Gly Ala Leu Ile Gly Ser Ser Arg Asn Gln Arg Phe
850 855 860
Glu Lys Leu Lys Ala Ala Ile His Asp Leu Ala Ser Ser Glu Ser Phe
865 870 875 880
Ile Ile Met Leu Leu Asn Asn Thr Arg Leu Asp Asn Ile Ser Asp Asn
885 890 895
Leu Ala Lys Ile Asp Gly Ile Leu Thr Asn Asn Thr Arg Phe Leu Ser
900 905 910
Asn Ala Thr Val Ser Lys Thr Leu Gln Thr Leu Gly Gly Ser Leu Ile
915 920 925
Arg Glu Cys Val Glu Ala Leu Asn Lys Arg Ser Pro Ser Ser Leu Asn
930 935 940
Asn Ala Arg Leu Leu Ala Val Gln Thr Ile Leu Gly His Ala Ser Val
945 950 955 960
Pro Asp His Glu Thr Leu Thr Arg Ile Val Ser Gly Val Ala Ser Ala
965 970 975
Gln Lys Glu Ser Ala Gly Asp Asp Pro Asp Arg Trp Thr Arg Val Thr
980 985 990
Gly His Leu Asn Glu Leu Lys Leu Val Thr Thr Gln Ser Arg Val Asp
995 1000 1005
Lys Ala Thr Arg Arg Lys Leu Leu Met Ile Ile Thr Arg Asp Leu Lys
1010 1015 1020

Glu Ala Glu Val Ser Gln Glu Thr Val Leu Glu Thr Arg Trp Gln Glu
1025 1030 1035 1040
Asn Val Leu Lys Phe Gln Pro Ser Thr Ser Lys Glu Ile Glu Asp Phe
1045 1050 1055
Leu Gln Ser Ala Pro Ser Ala Lys Ala Arg Lys Phe Ala Glu Lys His
1060 1065 1070
Leu Arg Thr Leu Ile Thr Gln Phe Asn Gly His Glu Arg Pro Pro Ser
1075 1080 1085
Glu Ala Thr Ala Val Pro Met Asp Tyr Thr Pro Thr Pro Ile Pro Thr
1090 1095 1100
Pro Gln Ala Val Ser Thr Ala Thr Ala Glu Lys Gly Lys Ala Ala Trp
1105 1110 1115 1120
Asn Lys Ile Gln Gln Ala Phe Gln Asp Phe Asn Phe His Leu Ile Asp
1125 1130 1135
Ala Ser Asp Trp Gln Glu Met Ala Ser Glu Tyr Ser Arg His Gly Ser
1140 1145 1150
Ser Leu Pro Gly Thr Val Gly Pro Lys Leu Val Arg Phe Met Glu Ser
1155 1160 1165
Ile Ser Asn Thr Leu Asp Asp Ile Leu Thr Gln Lys Leu Ala Ser Leu
1170 1175 1180
Leu Pro Asn Gly Pro Ala Phe Arg Pro Pro Ala Phe Asp Trp Ile Ala
1185 1190 1195 1200
Pro Tyr Gln Thr Arg Val Asn Ala Phe Leu Lys Thr Ile Gly Leu Pro
1205 1210 1215
Met Val Arg Asn Leu Ala Asp Lys Ile His His Gln Cys Gln Thr Val
1220 1225 1230
Ser His Ala Val Gln Ser Ala Asp Leu Gln Gln Ala Thr Val Gly Thr
1235 1240 1245
Ser Leu Glu Arg Pro Ala Ala Glu Tyr Cys Arg Ile Leu Ser Asp Met
1250 1255 1260
Gln Val Ala Phe Asn Asp His Gly Ile Ala Val Arg Ser Glu Ala Ala
1265 1270 1275 1280
Ala Tyr Thr Asp Ala Ile Asn Ser Pro Ala Asn Val Val Thr Pro Pro
1285 1290 1295
Lys Pro Asn Leu Glu Ala Pro Lys Lys Leu Ile Thr Ala Thr Asp Ala
1300 1305 1310
Leu Thr Val Glu Asp Phe Pro Asp Phe Leu Lys Thr Ser Ile Leu Gln
1315 1320 1325
Gln Glu Gln Arg Leu Ile Ala Leu Gln Arg Ala Glu Phe Gln Gln Leu
1330 1335 1340
Glu Ala Ser Ile Ser Ala Ala Glu Arg Leu Arg Gln Ser Thr Arg Asp
1345 1350 1355 1360
Glu Ile Ala Gly Lys Met Ala Thr Ala Ile Thr Gln Leu Leu Pro Arg
1365 1370 1375
Ala Pro Val Ala Ile Ser Ser Arg Pro Leu Asn Leu Ser Lys Pro Ile
1380 1385 1390
Asp Phe Leu Ser Ser Thr Val Tyr Asp Lys Ile Leu Asp Lys Glu Pro
1395 1400 1405
Tyr Glu Thr Ala Ile Ala Gly Phe Ala Trp Leu Glu Ile Ala Thr Lys
1410 1415 1420
Ser Val Met Val Tyr Ser Gln Gln Asn Glu Thr Gln Gln Leu Asn Val
1425 1430 1435 1440
Leu Leu Ser Glu Val Glu Lys Gln Ser Thr Val Ala Gln Arg Leu His
1445 1450 1455
Asp Leu Glu Leu Ser Ala Lys Asn Thr Asp Asp Val Lys Val Leu Lys
1460 1465 1470
Gln Ala Leu Asp Glu Leu Ala Pro Leu Arg Val Lys Gly Gly Lys Thr
1475 1480 1485
Thr Val Asp Ala Trp Lys Gln Lys Leu Glu Ser Ile Glu Ser Leu Leu
1490 1495 1500
Arg Ala Thr Arg Thr Ala Gly Glu Ile Ser Ser Glu Leu Glu Arg Ile

1505 1510 1515 1520
Gly Thr Gln Ala Val Gly Thr Ile Thr Val Arg Asp Leu Gly Thr Leu
1525 1530 1535
Ser Asp Gln Cys Arg Glu Ala Ala Asn Phe Leu Arg Gln Ala Ser Leu
1540 1545 1550
Pro Glu Gly Phe Ser Asp Ile Gly Thr Lys Leu Ser Glu Leu Gln Ala
1555 1560 1565
Tyr Ile Lys Tyr Lys Lys Gln Phe Leu Glu His Phe Glu Thr Thr Gln
1570 1575 1580
Pro Asn Val Phe Gln Arg Phe Pro Leu Ser Gln Asn Ile Thr Glu Asn
1585 1590 1595 1600
Val Pro Ala Arg Pro Ala Met Asp Ser Val Ala Arg Leu Thr Asn His
1605 1610 1615
Leu His Val Arg Gly Ser Ala Pro His Phe Thr Thr Trp Ile Glu Thr
1620 1625 1630
Leu Pro Thr Val Asp Pro Glu Lys Pro Thr His Val Pro Ala His Gly
1635 1640 1645
Gly Ala Pro Leu His Arg Gln Ile Thr Tyr Ser Asn Val Leu Glu Ala
1650 1655 1660
Leu Phe Ser Leu Cys Ser Thr Thr Leu Thr Pro Val Pro Thr Ala Pro
1665 1670 1675 1680
Gly Leu Glu Ile Ala Thr Arg Ala Arg Arg Gly Ala Glu Ala Ala Thr
1685 1690 1695
Trp Met Asp Arg Gln Trp Pro Asp Ile Ala Gln Thr Leu Gln Asp Val
1700 1705 1710
Leu Asp Thr Tyr Glu His Thr Thr Ala His Ala Asn Arg Asp Ala Ala
1715 1720 1725
Phe Asn Thr Phe Leu Ala Met Cys Val Phe Thr Gln Ile Ile Arg Gly
1730 1735 1740
Ala Ser Arg Ala Val Thr Leu Pro Lys Leu Pro Ser Thr Ala Val Asp
1745 1750 1755 1760
Phe Pro Glu Glu Ile Val Leu Thr Pro Arg Glu Cys Thr Thr Leu Val
1765 1770 1775
Thr Ala Met Trp Pro Thr Leu Ala Ala Ile Leu Arg Leu Lys Ser
1780 1785 1790
Tyr Ser Glu Ala Leu Gly Leu Met Ser Arg Phe Leu Pro Leu Met Phe
1795 1800 1805
Gln Ala Leu Pro His Leu Thr Leu Glu Ala Gln Val Lys Asn Gly Pro
1810 1815 1820
His Asn Thr Pro Pro Gln Leu Arg Cys Phe Ala Lys Thr Glu Ala Ile
1825 1830 1835 1840
Pro Tyr Phe Pro Ala Gln Trp Gln Ser Ala Asn Leu Glu Gln Ser Leu
1845 1850 1855
Trp Gly Gln Thr Asp Phe Leu Gln Ile Cys Asp Asn Asn Gln Arg Lys
1860 1865 1870
Ala Arg Val Ala Ala Val Thr Trp Ala Leu Thr Thr Ile Asp Gly Val
1875 1880 1885
Val Leu Asp Gln Leu Trp Ser Thr Phe Lys Pro Met Thr Ala Ala Ser
1890 1895 1900
Asp Asp Thr Tyr Val Asp Leu Val Glu Thr Leu His Leu Thr Thr Phe
1905 1910 1915 1920
Gly Pro Arg Gly Pro Thr Pro Arg Arg Glu Thr Thr Thr Glu His Pro
1925 1930 1935
Pro Tyr Glu Tyr Gly Gln Pro Thr Gly Tyr Cys Ile Ser Gly Gln Ser
1940 1945 1950
Thr Thr Pro Val Gln Ala Ser Asn Thr Pro Val Ser Ala Phe Glu Ala
1955 1960 1965
Val Leu Gly Ala Met Val Phe His Val Pro Ile Arg Ile Phe Leu Ala
1970 1975 1980
Ala Thr Pro Lys Arg Leu Gly Gln Ala Arg Gly Gly Met Gly Leu Leu
1985 1990 1995 2000

Thr Pro Ile Leu Glu Cys Val Pro Asp Val Glu Pro Phe Lys Ser Leu
2005 2010 2015
Tyr Asn Ala Pro Arg Lys Pro Val Pro Ile Glu Thr Leu Pro Ala Ser
2020 2025 2030
Leu His Pro His Asp Glu Arg Gln Val Phe Leu Arg Gln Ala Gln Trp
2035 2040 2045
Leu Ser Tyr Arg Phe Thr Pro His Glu Ala Ala Arg Ser Ser Thr Pro
2050 2055 2060
Pro Leu Leu Val Val Ile Asp Pro Glu Asn Leu Val Thr Ala Thr Tyr
2065 2070 2075 2080
Ser Ser Gly Gly Pro Ala Asn Phe Glu Ser Arg Pro Phe Tyr Val Met
2085 2090 2095
Pro Gly Pro Tyr Pro Pro Asp Trp Pro Lys Thr Leu Ser Val Thr Ser
2100 2105 2110
Asn Thr Ser Val Thr His Leu Ser His Asp Glu Ile Cys Asn Leu Phe
2115 2120 2125
Thr Thr Leu Ser Arg Glu His Gly Thr Val Gln Gly Arg Asp Ile Phe
2130 2135 2140
Ala Ala Ala Pro Thr Asn Val Thr Pro Glu Gln Thr Ala Asn Pro Pro
2145 2150 2155 2160
Ala Trp Glu Thr Asp Asn Arg Leu Ile Thr Gln Thr Glu Thr Ala Lys
2165 2170 2175
Lys Pro His Ile Ile Pro Ala Ser Pro Lys Ala Arg Thr Asp Pro Pro
2180 2185 2190
Val Glu Thr Thr Thr His His Ser Gln Gly Gln Ala Ser Gln His Ala
2195 2200 2205
Asn Ser Asn Val Asn Gln Pro Gly Gln Ile Thr Ser His Ala Ser Arg
2210 2215 2220
Asn Thr Pro Ser Thr Ala Pro Gln Ala Ser Ser Ser Pro Glu Lys Phe
2225 2230 2235 2240
Asn Thr Gln Thr Val Pro Arg Leu Ile Ser Gln Thr Ser Glu Thr Ala
2245 2250 2255
His Ile Asn Gln Pro Ala Ser Gly Gln Val Thr Glu Pro Lys Gly Ile
2260 2265 2270
Phe Gly Thr Tyr Lys Pro Arg Val Leu Thr Glu Pro Ala Lys Pro Ala
2275 2280 2285
Asn Ala Gly Val Ala Ser Arg Gln Pro Glu Ala Thr Thr Thr Val Pro
2290 2295 2300
Lys Leu Pro Ile Asn Pro Pro Thr Ala Arg Val Phe Ile Gly Thr Ala
2305 2310 2315 2320
Ser Lys Leu Ser Pro Ala Val Glu Glu Ser His Gly Ala Thr Pro Asp
2325 2330 2335
Ala His Gln Ser Lys Ile Asp Arg Glu Lys Tyr Ala Glu Ser Arg Pro
2340 2345 2350
Arg Arg Thr Pro His Leu Glu Glu Gly Pro Arg Glu Pro His Val Asn
2355 2360 2365
Thr Pro Thr Ser Ala His Ile Asn Val Pro Ser Ser Gln Gly Gln Lys
2370 2375 2380
Thr Val His Gly Arg Glu Asn Pro Gly Leu Gln Thr Ala Thr Pro Ser
2385 2390 2395 2400
Ala Pro Gln Pro Thr Ala Ser Asn Pro Arg Ile Gln Tyr Thr Leu Pro
2405 2410 2415
Arg Thr Asp Gly Arg Leu Leu His Asp Glu Ser Glu Val Glu Ser Thr
2420 2425 2430
Pro Thr Glu Glu Val Lys Arg Ser Pro Lys Thr Gln Asp Val Ser His
2435 2440 2445
Gly Pro Glu Pro Asp Asp Ser Arg Trp Thr Ala Pro Leu Gly Pro Thr
2450 2455 2460
Ile Glu Ile His Arg Leu Glu His Pro Gln Ile Leu Lys Asn Ile Thr
2465 2470 2475 2480
Ser Leu Thr Val Pro Thr Pro Arg Val Thr Pro Ile Pro Pro Thr Asn

2485 2490 2495
Ile Trp Ile Pro Leu Ser His Val Asn Ile Gln His Glu Glu Ile Thr
2500 2505 2510
Arg Ala Lys Asn Val Leu Met Arg Phe Ile Gln Asn Val Arg Arg Lys
2515 2520 2525
Leu Gln Ala Ser Ser Asp Ala Leu Ser Glu Ala Ile Ala Arg Ile Lys
2530 2535 2540
Phe Leu Tyr Leu
2545

<210> 144
<211> 510
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(510)

<400> 144
atg tcg tcc ttg cgg gtt aag gag cca atc gtt cag gga cgc ctt gag 48
Met Ser Ser Leu Arg Val Lys Glu Pro Ile Val Gln Gly Arg Leu Glu
1 5 10 15

cac gat tac cca aat cac ccg ctg gtt gcc gag atg aac aac ctt ccc 96
His Asp Tyr Pro Asn His Pro Leu Val Ala Glu Met Asn Asn Leu Pro
20 25 30

cag ggt gac atg tct ccc gcc cag tat gcc atc gcg aaa cgt aat tat 144
Gln Gly Asp Met Ser Pro Ala Gln Tyr Ala Ile Ala Lys Arg Asn Tyr
35 40 45

ctg gtg ttt tta acg gcc aaa cat cat tac gac atg tat atg caa aag 192
Leu Val Phe Leu Thr Ala Lys His His Tyr Asp Met Tyr Met Gln Lys
50 55 60

aag aat gga att ctg cgc aaa gac cac ctc cgg ggc ctt cgc ggc aaa 240
Lys Asn Gly Ile Leu Arg Lys Asp His Leu Arg Gly Leu Arg Gly Lys
65 70 75 80

aag gac gct agt tct agt atc tcg ggc gtt ttg tcc ggg tcc ggc tcg 288
Lys Asp Ala Ser Ser Ser Ile Ser Gly Val Leu Ser Gly Ser Gly Ser
85 90 95

gcc gcc ccg agc gtt gcc ccg gtg gcc tcc acc ctc ggg tca aat agc 336
Ala Ala Pro Ser Val Ala Pro Val Ala Ser Thr Leu Gly Ser Asn Ser
100 105 110

ttt act acg atc tcc agc ggg ccc cat tca ttg ata ggc tcg atg ggc 384
Phe Thr Thr Ile Ser Ser Gly Pro His Ser Leu Ile Gly Ser Met Gly
115 120 125

ccc gcg ccc ggt ggc ggc gga ccc ggg agc gtg gcg tct tct ggc ata 432
Pro Ala Pro Gly Gly Gly Gly Pro Gly Ser Val Ala Ser Ser Gly Ile
130 135 140

ggg tcc act tcg tta tca cct agc gac gca act act ctg gac aca aga 480
Gly Ser Thr Ser Leu Ser Pro Ser Asp Ala Thr Thr Leu Asp Thr Arg
145 150 155 160

cgc tcg tct caa aat aaa aaa agc aag tga 510
Arg Ser Ser Gln Asn Lys Lys Ser Lys
165 170

<210> 145
<211> 169
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 145
Met Ser Ser Leu Arg Val Lys Glu Pro Ile Val Gln Gly Arg Leu Glu
1 5 10 15
His Asp Tyr Pro Asn His Pro Leu Val Ala Glu Met Asn Asn Leu Pro
20 25 30
Gln Gly Asp Met Ser Pro Ala Gln Tyr Ala Ile Ala Lys Arg Asn Tyr
35 40 45
Leu Val Phe Leu Thr Ala Lys His His Tyr Asp Met Tyr Met Gln Lys
50 55 60
Lys Asn Gly Ile Leu Arg Lys Asp His Leu Arg Gly Leu Arg Gly Lys
65 70 75 80
Lys Asp Ala Ser Ser Ser Ile Ser Gly Val Leu Ser Gly Ser Gly Ser
85 90 95
Ala Ala Pro Ser Val Ala Pro Val Ala Ser Thr Leu Gly Ser Asn Ser
100 105 110
Phe Thr Thr Ile Ser Ser Gly Pro His Ser Leu Ile Gly Ser Met Gly
115 120 125
Pro Ala Pro Gly Gly Gly Gly Pro Gly Ser Val Ala Ser Ser Gly Ile
130 135 140
Gly Ser Thr Ser Leu Ser Pro Ser Asp Ala Thr Thr Leu Asp Thr Arg
145 150 155 160
Arg Ser Ser Gln Asn Lys Lys Ser Lys
165

<210> 146
<211> 1347
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1347)

<400> 146
atg gcc tcg ggc cgc ttg cct aac ctg gct gaa gac gaa gct gcc tgt 48
Met Ala Ser Gly Arg Leu Pro Asn Leu Ala Glu Asp Glu Ala Ala Cys
1 5 10 15

cat ggg cgc ggt tct tat cct gcc cat cgt tgg ctg gat ggt tct cgg 96
His Gly Arg Gly Ser Tyr Pro Ala His Arg Trp Leu Asp Gly Ser Arg
20 25 30

ctg ggc tta gat ctc gcg gcc tct ata cgc tca atc gga cta tgc ccc 144
Leu Gly Leu Asp Leu Ala Ala Ser Ile Arg Ser Ile Gly Leu Cys Pro
35 40 45

gaa tgc tac gtg tgt ttt gtg acg tac ggg ctc ggt gcc tgg gac gga 192
Glu Cys Tyr Val Cys Phe Val Thr Tyr Gly Leu Gly Ala Trp Asp Gly

```
     50                  55                  60

cga ccc cca aaa tgg gct tgc acg tta att tct gcc cca tcc ttt cag    240
Arg Pro Pro Lys Trp Ala Cys Thr Leu Ile Ser Ala Pro Ser Phe Gln
 65                 70                  75                  80

acc gcc ctg aac gag att gct acc ggc tgg agg ccg gat aac cca cct    288
Thr Ala Leu Asn Glu Ile Ala Thr Gly Trp Arg Pro Asp Asn Pro Pro
                 85                  90                  95

aaa aac ggc gac gtg aga tcg cgg ctg cac gat atc ggc cgg tcg ttg    336
Lys Asn Gly Asp Val Arg Ser Arg Leu His Asp Ile Gly Arg Ser Leu
            100                 105                 110

tta gag gcg tac gcg tgg gtg ctg agg tgt atc tgc acc ggc gtg ggc    384
Leu Glu Ala Tyr Ala Trp Val Leu Arg Cys Ile Cys Thr Gly Val Gly
        115                 120                 125

tgt ccc agc gac gaa ggt tta agt tta acg gcc gtt ccc agg agc gcg    432
Cys Pro Ser Asp Glu Gly Leu Ser Leu Thr Ala Val Pro Arg Ser Ala
    130                 135                 140

tgg tcc agg tat ctt gtt gtg tcg ttt cag agg gcg tgt tgt ctc gtg    480
Trp Ser Arg Tyr Leu Val Val Ser Phe Gln Arg Ala Cys Cys Leu Val
145                 150                 155                 160

tgc aaa acc ctt aac tgt cgc cag cgg ttt ccc ctg gtt acc tgt ttg    528
Cys Lys Thr Leu Asn Cys Arg Gln Arg Phe Pro Leu Val Thr Cys Leu
                165                 170                 175

cca cag cac gcg ttg gat ctg ccc gtt ttg cgg aag aaa tgg aac ggg    576
Pro Gln His Ala Leu Asp Leu Pro Val Leu Arg Lys Lys Trp Asn Gly
            180                 185                 190

ggc ggt tgt gtc tcg atg cag ttg aac gtt ccg tcc att tcc aga cgt    624
Gly Gly Cys Val Ser Met Gln Leu Asn Val Pro Ser Ile Ser Arg Arg
        195                 200                 205

ctg gga gct aac ctg aat gag agt gtc ccg ggc cct tcg gac gcg gga    672
Leu Gly Ala Asn Leu Asn Glu Ser Val Pro Gly Pro Ser Asp Ala Gly
    210                 215                 220

ctt ctg gcc tcg ctg cga gaa ctg gcg ccg acc gtc ccg tgc ggc aac    720
Leu Leu Ala Ser Leu Arg Glu Leu Ala Pro Thr Val Pro Cys Gly Asn
225                 230                 235                 240

ccg ttt aac gcc ctt ctg agg agc ctg acg ttt agg gcc ctg ctg agc    768
Pro Phe Asn Ala Leu Leu Arg Ser Leu Thr Phe Arg Ala Leu Leu Ser
                245                 250                 255

atg tcc cgt gtt gtg ttg ccc ata ggg gag tcg acc gag acg gag ata    816
Met Ser Arg Val Val Leu Pro Ile Gly Glu Ser Thr Glu Thr Glu Ile
            260                 265                 270

tcc cgc gac ctg ggt caa aaa gta ttg gcg tat aac gta ctg ttt ccg    864
Ser Arg Asp Leu Gly Gln Lys Val Leu Ala Tyr Asn Val Leu Phe Pro
        275                 280                 285

tgt atc tct cta cct gtt tgg tcc cag gtg gtg gcc agg agc gta ctg    912
Cys Ile Ser Leu Pro Val Trp Ser Gln Val Val Ala Arg Ser Val Leu
    290                 295                 300
```

gag aag acc gtg cca gct ccg agg gtg gtg gtg tgc ctc gag tgc ggt 960
Glu Lys Thr Val Pro Ala Pro Arg Val Val Val Cys Leu Glu Cys Gly
305 310 315 320

tac tgt ttg aat ttc gga cgg ggt aaa ttt gaa acc gta aat ttc cca 1008
Tyr Cys Leu Asn Phe Gly Arg Gly Lys Phe Glu Thr Val Asn Phe Pro
325 330 335

ccc acc aac gtg ttt ttc agt cgc gac caa aaa gag aag cag ctt tcg 1056
Pro Thr Asn Val Phe Phe Ser Arg Asp Gln Lys Glu Lys Gln Leu Ser
340 345 350

atc tgt gct acc acc ggg cgc gta tat tgc tct tat tgt ggc gga tcg 1104
Ile Cys Ala Thr Thr Gly Arg Val Tyr Cys Ser Tyr Cys Gly Gly Ser
355 360 365

cac atg cgg gtt ata tct cta ttc gag atc acg tgt gtt gga gat ccc 1152
His Met Arg Val Ile Ser Leu Phe Glu Ile Thr Cys Val Gly Asp Pro
370 375 380

tat tta cgg tgc gtg ctc gcc aac aac gcg gca cat gcc ata cga gac 1200
Tyr Leu Arg Cys Val Leu Ala Asn Asn Ala Ala His Ala Ile Arg Asp
385 390 395 400

gcg aac tcc ctg gtt agc gtc gtc gtg ccc tgt ttg gcg tcg ccg gac 1248
Ala Asn Ser Leu Val Ser Val Val Val Pro Cys Leu Ala Ser Pro Asp
405 410 415

tgc gcg acc ggc cta tta aag cat ttg cgt gtg gcc gag ctg ttt tat 1296
Cys Ala Thr Gly Leu Leu Lys His Leu Arg Val Ala Glu Leu Phe Tyr
420 425 430

ttg acc tcg tct ata tcg tcc ctg tcc tgt gga aag tgt aat aga agt 1344
Leu Thr Ser Ser Ile Ser Ser Leu Ser Cys Gly Lys Cys Asn Arg Ser
435 440 445

taa 1347

<210> 147
<211> 448
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 147
Met Ala Ser Gly Arg Leu Pro Asn Leu Ala Glu Asp Glu Ala Ala Cys
1 5 10 15
His Gly Arg Gly Ser Tyr Pro Ala His Arg Trp Leu Asp Gly Ser Arg
20 25 30
Leu Gly Leu Asp Leu Ala Ala Ser Ile Arg Ser Ile Gly Leu Cys Pro
35 40 45
Glu Cys Tyr Val Cys Phe Val Thr Tyr Gly Leu Gly Ala Trp Asp Gly
50 55 60
Arg Pro Pro Lys Trp Ala Cys Thr Leu Ile Ser Ala Pro Ser Phe Gln
65 70 75 80
Thr Ala Leu Asn Glu Ile Ala Thr Gly Trp Arg Pro Asp Asn Pro Pro
85 90 95
Lys Asn Gly Asp Val Arg Ser Arg Leu His Asp Ile Gly Arg Ser Leu
100 105 110

Leu Glu Ala Tyr Ala Trp Val Leu Arg Cys Ile Cys Thr Gly Val Gly
115 120 125

Cys Pro Ser Asp Glu Gly Leu Ser Leu Thr Ala Val Pro Arg Ser Ala
130 135 140

Trp Ser Arg Tyr Leu Val Val Ser Phe Gln Arg Ala Cys Cys Leu Val
145 150 155 160

Cys Lys Thr Leu Asn Cys Arg Gln Arg Phe Pro Leu Val Thr Cys Leu
165 170 175

Pro Gln His Ala Leu Asp Leu Pro Val Leu Arg Lys Lys Trp Asn Gly
180 185 190

Gly Gly Cys Val Ser Met Gln Leu Asn Val Pro Ser Ile Ser Arg Arg
195 200 205

Leu Gly Ala Asn Leu Asn Glu Ser Val Pro Gly Pro Ser Asp Ala Gly
210 215 220

Leu Leu Ala Ser Leu Arg Glu Leu Ala Pro Thr Val Pro Cys Gly Asn
225 230 235 240

Pro Phe Asn Ala Leu Leu Arg Ser Leu Thr Phe Arg Ala Leu Leu Ser
245 250 255

Met Ser Arg Val Val Leu Pro Ile Gly Glu Ser Thr Glu Thr Glu Ile
260 265 270

Ser Arg Asp Leu Gly Gln Lys Val Leu Ala Tyr Asn Val Leu Phe Pro
275 280 285

Cys Ile Ser Leu Pro Val Trp Ser Gln Val Val Ala Arg Ser Val Leu
290 295 300

Glu Lys Thr Val Pro Ala Pro Arg Val Val Val Cys Leu Glu Cys Gly
305 310 315 320

Tyr Cys Leu Asn Phe Gly Arg Gly Lys Phe Glu Thr Val Asn Phe Pro
325 330 335

Pro Thr Asn Val Phe Phe Ser Arg Asp Gln Lys Glu Lys Gln Leu Ser
340 345 350

Ile Cys Ala Thr Thr Gly Arg Val Tyr Cys Ser Tyr Cys Gly Gly Ser
355 360 365

His Met Arg Val Ile Ser Leu Phe Glu Ile Thr Cys Val Gly Asp Pro
370 375 380

Tyr Leu Arg Cys Val Leu Ala Asn Asn Ala Ala His Ala Ile Arg Asp
385 390 395 400

Ala Asn Ser Leu Val Ser Val Val Val Pro Cys Leu Ala Ser Pro Asp
405 410 415

Cys Ala Thr Gly Leu Leu Lys His Leu Arg Val Ala Glu Leu Phe Tyr
420 425 430

Leu Thr Ser Ser Ile Ser Ser Leu Ser Cys Gly Lys Cys Asn Arg Ser
435 440 445

<210> 148
<211> 675
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(675)

<400> 148
atg agt tcc gga aaa cgt tta gtt gac gag ctg tgc gat tta gtt gtc 48
Met Ser Ser Gly Lys Arg Leu Val Asp Glu Leu Cys Asp Leu Val Val
1 5 10 15

tct tac ctg gga ccc tct gga atc tcc ctg gac ttg gag agg tgt cag 96
Ser Tyr Leu Gly Pro Ser Gly Ile Ser Leu Asp Leu Glu Arg Cys Gln

20 25 30

gac ggg gct ccc gta tac gct aaa ggc ggg gcg gtc ccc gtg tgc acc 144
Asp Gly Ala Pro Val Tyr Ala Lys Gly Gly Ala Val Pro Val Cys Thr
35 40 45

gtg cgc ctg cag cac ggc tgc gtc tat cat ctc gag ttt gtg tat aag 192
Val Arg Leu Gln His Gly Cys Val Tyr His Leu Glu Phe Val Tyr Lys
50 55 60

ttt tgg ctc cac aaa cta gag aga ctg gcc tac ccg ttt gcc ccg tgt 240
Phe Trp Leu His Lys Leu Glu Arg Leu Ala Tyr Pro Phe Ala Pro Cys
65 70 75 80

ttt gta att atc aac aac ggt ttg gcc acc acg ctg aaa tgt ttt ttg 288
Phe Val Ile Ile Asn Asn Gly Leu Ala Thr Thr Leu Lys Cys Phe Leu
85 90 95

tgt aag cca cgt gac gcc gat gcc cag ttt gga aaa aac ctg cct ata 336
Cys Lys Pro Arg Asp Ala Asp Ala Gln Phe Gly Lys Asn Leu Pro Ile
100 105 110

aat tcg gac gtg tat ctt gag agg aac tcg tcc gtg tcc ctg ggc cag 384
Asn Ser Asp Val Tyr Leu Glu Arg Asn Ser Ser Val Ser Leu Gly Gln
115 120 125

gac gat ttt atg aaa ttt aag gca cgt ctg gtt ttc tcc gga gac cta 432
Asp Asp Phe Met Lys Phe Lys Ala Arg Leu Val Phe Ser Gly Asp Leu
130 135 140

aac gtt tac agc tcc atg gtc ata tgc cgc acc tac ttt acg gag cac 480
Asn Val Tyr Ser Ser Met Val Ile Cys Arg Thr Tyr Phe Thr Glu His
145 150 155 160

cga cag gtt tta cag ttt ttg gtc gtg act cca aag agc gct aaa cgg 528
Arg Gln Val Leu Gln Phe Leu Val Val Thr Pro Lys Ser Ala Lys Arg
165 170 175

tta aaa acc ctt ctt aga acg gtt ttt gcc ctg acg ggt cac tcc gac 576
Leu Lys Thr Leu Leu Arg Thr Val Phe Ala Leu Thr Gly His Ser Asp
180 185 190

ggc ctc ggt gcg ttg agg cga acg ggc tcc gtg gcc cgc cct tcg ggg 624
Gly Leu Gly Ala Leu Arg Arg Thr Gly Ser Val Ala Arg Pro Ser Gly
195 200 205

tcg gag ttg aag gat att ggg cgc gga gag cgt gcg gcg atg acc aat 672
Ser Glu Leu Lys Asp Ile Gly Arg Gly Glu Arg Ala Ala Met Thr Asn
210 215 220

taa 675

225

<210> 149
<211> 224
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 149

Met Ser Ser Gly Lys Arg Leu Val Asp Glu Leu Cys Asp Leu Val Val
1   5   10   15
Ser Tyr Leu Gly Pro Ser Gly Ile Ser Leu Asp Leu Glu Arg Cys Gln
            20   25   30
Asp Gly Ala Pro Val Tyr Ala Lys Gly Gly Ala Val Pro Val Cys Thr
        35   40   45
Val Arg Leu Gln His Gly Cys Val Tyr His Leu Glu Phe Val Tyr Lys
    50   55   60
Phe Trp Leu His Lys Leu Glu Arg Leu Ala Tyr Pro Phe Ala Pro Cys
65   70   75   80
Phe Val Ile Ile Asn Asn Gly Leu Ala Thr Thr Leu Lys Cys Phe Leu
                85   90   95
Cys Lys Pro Arg Asp Ala Asp Ala Gln Phe Gly Lys Asn Leu Pro Ile
            100   105   110
Asn Ser Asp Val Tyr Leu Glu Arg Asn Ser Ser Val Ser Leu Gly Gln
        115   120   125
Asp Asp Phe Met Lys Phe Lys Ala Arg Leu Val Phe Ser Gly Asp Leu
    130   135   140
Asn Val Tyr Ser Ser Met Val Ile Cys Arg Thr Tyr Phe Thr Glu His
145   150   155   160
Arg Gln Val Leu Gln Phe Leu Val Val Thr Pro Lys Ser Ala Lys Arg
                165   170   175
Leu Lys Thr Leu Leu Arg Thr Val Phe Ala Leu Thr Gly His Ser Asp
            180   185   190
Gly Leu Gly Ala Leu Arg Arg Thr Gly Ser Val Ala Arg Pro Ser Gly
        195   200   205
Ser Glu Leu Lys Asp Ile Gly Arg Gly Glu Arg Ala Ala Met Thr Asn
    210   215   220

<210> 150
<211> 1374
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1374)

<400> 150
atg ttc gtc ccg tgg caa ctg gag acc ctc atg aga cac tgg ccc tct   48
Met Phe Val Pro Trp Gln Leu Glu Thr Leu Met Arg His Trp Pro Ser
1   5   10   15

ctg cgg gga ctc gta gaa caa tcc ttc ctc ccc ggt acc ccg gac gga   96
Leu Arg Gly Leu Val Glu Gln Ser Phe Leu Pro Gly Thr Pro Asp Gly
            20   25   30

gct ttt aac agc ccg gta tta atc cac act cag gac tct cta caa ccc   144
Ala Phe Asn Ser Pro Val Leu Ile His Thr Gln Asp Ser Leu Gln Pro
        35   40   45

gcc tca tcg tgc agg gtg tgt agc ctc ctg ttc act ctg gtc cgg aca   192
Ala Ser Ser Cys Arg Val Cys Ser Leu Leu Phe Thr Leu Val Arg Thr
    50   55   60

ttc cca ccc ccc gac tct ttc ttc gaa gac tac ggc tgg ttg tgc ctc   240
Phe Pro Pro Pro Asp Ser Phe Phe Glu Asp Tyr Gly Trp Leu Cys Leu
65   70   75   80

```
acc tgc cta tac gcc ccc cga tca tgg acg gct acc ctc atg gtg gct    288
Thr Cys Leu Tyr Ala Pro Arg Ser Trp Thr Ala Thr Leu Met Val Ala
                85                  90                  95

gcc gac ctt ttg gaa cta acg cac gtg tac ttc ccg caa tgc gtg aaa    336
Ala Asp Leu Leu Glu Leu Thr His Val Tyr Phe Pro Gln Cys Val Lys
            100                 105                 110

gat ggg cca gta tac acc gcc caa agc atc ctc gga atc gac gtc cag    384
Asp Gly Pro Val Tyr Thr Ala Gln Ser Ile Leu Gly Ile Asp Val Gln
        115                 120                 125

ctg cac ttc ttc gca acc cgc tgc ttc cga ccc atc gac aga gaa caa    432
Leu His Phe Phe Ala Thr Arg Cys Phe Arg Pro Ile Asp Arg Glu Gln
    130                 135                 140

ata ctc cac aca tct cat tta aat ttt tta caa acc gag ttt att agg    480
Ile Leu His Thr Ser His Leu Asn Phe Leu Gln Thr Glu Phe Ile Arg
145                 150                 155                 160

ggc atg tta gaa ggc acg att ccg gga tcg ttc tgt ttt aaa acg tcc    528
Gly Met Leu Glu Gly Thr Ile Pro Gly Ser Phe Cys Phe Lys Thr Ser
                165                 170                 175

tgg ccg cgc aca gaa aag gac gac caa caa cct acc gtt gcg tgt tgt    576
Trp Pro Arg Thr Glu Lys Asp Asp Gln Gln Pro Thr Val Ala Cys Cys
            180                 185                 190

tcc gtt ggc cgc gga agt cac acc aac cgg gat aac cgc cta ccc gag    624
Ser Val Gly Arg Gly Ser His Thr Asn Arg Asp Asn Arg Leu Pro Glu
        195                 200                 205

gac ctg gaa gag gcg ttc aac tcc acg aac gcc gag gaa aag ccc agc    672
Asp Leu Glu Glu Ala Phe Asn Ser Thr Asn Ala Glu Glu Lys Pro Ser
    210                 215                 220

ctc ctc ggc gtc ttt tcg gca acg tgg gca gaa tcc cag ctt ctt ggc    720
Leu Leu Gly Val Phe Ser Ala Thr Trp Ala Glu Ser Gln Leu Leu Gly
225                 230                 235                 240

tcc gac aca caa cag gca gat acc cat tta caa ccc tcc gcc ttc cca    768
Ser Asp Thr Gln Gln Ala Asp Thr His Leu Gln Pro Ser Ala Phe Pro
                245                 250                 255

acc cca gaa gat gct gac caa tca cag ggc ccc tgc ctg atg cac cca    816
Thr Pro Glu Asp Ala Asp Gln Ser Gln Gly Pro Cys Leu Met His Pro
            260                 265                 270

acg ctc aac cta aaa aca aaa aac cac acc gca tcc ata tgc gtt cta    864
Thr Leu Asn Leu Lys Thr Lys Asn His Thr Ala Ser Ile Cys Val Leu
        275                 280                 285

tgc gag tgt ctg gcc gca cac ccg gac gcc ggt ccg gtt ctg aaa gat    912
Cys Glu Cys Leu Ala Ala His Pro Asp Ala Gly Pro Val Leu Lys Asp
    290                 295                 300

ctg cgt cgc gac att ctg gaa aac atg gaa aac aac gtt aag ctc gtc    960
Leu Arg Arg Asp Ile Leu Glu Asn Met Glu Asn Asn Val Lys Leu Val
305                 310                 315                 320

aat cgc ata tcg tac atc cta aac gat ccg gac tca ctg tca cac gtg    1008
```

Asn Arg Ile Ser Tyr Ile Leu Asn Asp Pro Asp Ser Leu Ser His Val
325 330 335

cgc gac gaa cat ctg cgc ggc cta att aaa cgg tgc tcg gca caa gaa 1056
Arg Asp Glu His Leu Arg Gly Leu Ile Lys Arg Cys Ser Ala Gln Glu
340 345 350

atc cac aag cat ttt ttt tgc gac ccg gtg tgc gtc ctg aac acg tac 1104
Ile His Lys His Phe Phe Cys Asp Pro Val Cys Val Leu Asn Thr Tyr
355 360 365

tcg cac tgt ccc gcg gtt tta ttt aaa tgc cca cct ccc gaa aag tat 1152
Ser His Cys Pro Ala Val Leu Phe Lys Cys Pro Pro Pro Glu Lys Tyr
370 375 380

aag aag ctc aaa gct cgt ctg gca acc gga gag ttc cta gac tgc aac 1200
Lys Lys Leu Lys Ala Arg Leu Ala Thr Gly Glu Phe Leu Asp Cys Asn
385 390 395 400

aga ata ttt gac tgc gag acc tta cag acc ctg gcc gtc ctc ttt aag 1248
Arg Ile Phe Asp Cys Glu Thr Leu Gln Thr Leu Ala Val Leu Phe Lys
405 410 415

ggg tct caa ctg gcc aaa atc ggc aaa acc acg tcg ctc gag ata atc 1296
Gly Ser Gln Leu Ala Lys Ile Gly Lys Thr Thr Ser Leu Glu Ile Ile
420 425 430

cgt gaa ctc gga ttt caa ctg cgt cga cac aac att caa atc acc cac 1344
Arg Glu Leu Gly Phe Gln Leu Arg Arg His Asn Ile Gln Ile Thr His
435 440 445

ccg ttt caa acc tcc aac cta tac att taa 1374
Pro Phe Gln Thr Ser Asn Leu Tyr Ile
450 455

<210> 151
<211> 457
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 151
Met Phe Val Pro Trp Gln Leu Glu Thr Leu Met Arg His Trp Pro Ser
1 5 10 15
Leu Arg Gly Leu Val Glu Gln Ser Phe Leu Pro Gly Thr Pro Asp Gly
20 25 30
Ala Phe Asn Ser Pro Val Leu Ile His Thr Gln Asp Ser Leu Gln Pro
35 40 45
Ala Ser Ser Cys Arg Val Cys Ser Leu Leu Phe Thr Leu Val Arg Thr
50 55 60
Phe Pro Pro Pro Asp Ser Phe Phe Glu Asp Tyr Gly Trp Leu Cys Leu
65 70 75 80
Thr Cys Leu Tyr Ala Pro Arg Ser Trp Thr Ala Thr Leu Met Val Ala
85 90 95
Ala Asp Leu Leu Glu Leu Thr His Val Tyr Phe Pro Gln Cys Val Lys
100 105 110
Asp Gly Pro Val Tyr Thr Ala Gln Ser Ile Leu Gly Ile Asp Val Gln
115 120 125
Leu His Phe Phe Ala Thr Arg Cys Phe Arg Pro Ile Asp Arg Glu Gln
130 135 140
Ile Leu His Thr Ser His Leu Asn Phe Leu Gln Thr Glu Phe Ile Arg

```
145                 150                 155                 160
Gly Met Leu Glu Gly Thr Ile Pro Gly Ser Phe Cys Phe Lys Thr Ser
                165                 170                 175
Trp Pro Arg Thr Glu Lys Asp Asp Gln Gln Pro Thr Val Ala Cys Cys
            180                 185                 190
Ser Val Gly Arg Gly Ser His Thr Asn Arg Asp Asn Arg Leu Pro Glu
        195                 200                 205
Asp Leu Glu Glu Ala Phe Asn Ser Thr Asn Ala Glu Glu Lys Pro Ser
    210                 215                 220
Leu Leu Gly Val Phe Ser Ala Thr Trp Ala Glu Ser Gln Leu Leu Gly
225                 230                 235                 240
Ser Asp Thr Gln Gln Ala Asp Thr His Leu Gln Pro Ser Ala Phe Pro
                245                 250                 255
Thr Pro Glu Asp Ala Asp Gln Ser Gln Gly Pro Cys Leu Met His Pro
            260                 265                 270
Thr Leu Asn Leu Lys Thr Lys Asn His Thr Ala Ser Ile Cys Val Leu
        275                 280                 285
Cys Glu Cys Leu Ala Ala His Pro Asp Ala Gly Pro Val Leu Lys Asp
    290                 295                 300
Leu Arg Arg Asp Ile Leu Glu Asn Met Glu Asn Asn Val Lys Leu Val
305                 310                 315                 320
Asn Arg Ile Ser Tyr Ile Leu Asn Asp Pro Asp Ser Leu Ser His Val
                325                 330                 335
Arg Asp Glu His Leu Arg Gly Leu Ile Lys Arg Cys Ser Ala Gln Glu
            340                 345                 350
Ile His Lys His Phe Phe Cys Asp Pro Val Cys Val Leu Asn Thr Tyr
        355                 360                 365
Ser His Cys Pro Ala Val Leu Phe Lys Cys Pro Pro Glu Lys Tyr
    370                 375                 380
Lys Lys Leu Lys Ala Arg Leu Ala Thr Gly Glu Phe Leu Asp Cys Asn
385                 390                 395                 400
Arg Ile Phe Asp Cys Glu Thr Leu Gln Thr Leu Ala Val Leu Phe Lys
                405                 410                 415
Gly Ser Gln Leu Ala Lys Ile Gly Lys Thr Thr Ser Leu Glu Ile Ile
            420                 425                 430
Arg Glu Leu Gly Phe Gln Leu Arg Arg His Asn Ile Gln Ile Thr His
        435                 440                 445
Pro Phe Gln Thr Ser Asn Leu Tyr Ile
    450                 455
```

<210> 152
<211> 894
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(894)

<400> 152

```
atg cca aaa cag ccc aga agt cga ttg gcg tct cga gcg ccg tac gca      48
Met Pro Lys Gln Pro Arg Ser Arg Leu Ala Ser Arg Ala Pro Tyr Ala
  1               5                  10                  15

cct agc gtc agg cga ccg gac ggg ccc cag tcc acg cga ccg gca tcc      96
Pro Ser Val Arg Arg Pro Asp Gly Pro Gln Ser Thr Arg Pro Ala Ser
             20                  25                  30

agg cac ggc agc tgc aaa agc gaa atc atg cag tgg aaa aag tta gtt     144
```

```
Arg His Gly Ser Cys Lys Ser Glu Ile Met Gln Trp Lys Lys Leu Val
        35                  40                  45

tca gac acg cag ttt ttt tct gcc cta acg cgc cgc cac gag ctg ggg     192
Ser Asp Thr Gln Phe Phe Ser Ala Leu Thr Arg Arg His Glu Leu Gly
    50                  55                  60

gtg gac ttt tta aga gaa atg ggg acc ccg ata tgc acc tca aag tcc     240
Val Asp Phe Leu Arg Glu Met Gly Thr Pro Ile Cys Thr Ser Lys Ser
65                  70                  75                  80

gtt atg ttg ccg tta aac cta aaa acc atc gcc ccg ggt cgg tgc gtc     288
Val Met Leu Pro Leu Asn Leu Lys Thr Ile Ala Pro Gly Arg Cys Val
                85                  90                  95

tct ctc tca tca ttc gga cac tcg tca aac atg ggg ttc aac tgt tcg     336
Ser Leu Ser Ser Phe Gly His Ser Ser Asn Met Gly Phe Asn Cys Ser
            100                 105                 110

tcg tgc acg cca act gac agg tca gcg gtg tct ctg gac gca aac gcg     384
Ser Cys Thr Pro Thr Asp Arg Ser Ala Val Ser Leu Asp Ala Asn Ala
        115                 120                 125

ctc ggc gaa gat tcc gcc agg aaa aac agc gag ctg tgt tca gtg gcg     432
Leu Gly Glu Asp Ser Ala Arg Lys Asn Ser Glu Leu Cys Ser Val Ala
    130                 135                 140

tta acc ttt tac cac cac gcc gaa aag gtc gtg cag cac aag ggc ttt     480
Leu Thr Phe Tyr His His Ala Glu Lys Val Val Gln His Lys Gly Phe
145                 150                 155                 160

tac ctg tct ctg ctc agc cac tcc atg gaa gtc gtt agg aaa agc ttc     528
Tyr Leu Ser Leu Leu Ser His Ser Met Glu Val Val Arg Lys Ser Phe
                165                 170                 175

acg caa ccc ggg ttg ctc tac gcc cac cta gtc cta aaa acc ttt ggc     576
Thr Gln Pro Gly Leu Leu Tyr Ala His Leu Val Leu Lys Thr Phe Gly
            180                 185                 190

cac gat cct tta cct att ttt aca gtc gat gcc gat gag aga ctc gca     624
His Asp Pro Leu Pro Ile Phe Thr Val Asp Ala Asp Glu Arg Leu Ala
        195                 200                 205

ctc tgg gcg gtg ttc cac act aga gac cta cac ctg ggg gaa acc agt     672
Leu Trp Ala Val Phe His Thr Arg Asp Leu His Leu Gly Glu Thr Ser
    210                 215                 220

ctg cga ctc att atg gac aac ctt cca aat tat gac ata acg gtg gac     720
Leu Arg Leu Ile Met Asp Asn Leu Pro Asn Tyr Asp Ile Thr Val Asp
225                 230                 235                 240

tgc atc aag caa acg tac ata atg aag ttt aca ccc tcg cga ccg gac     768
Cys Ile Lys Gln Thr Tyr Ile Met Lys Phe Thr Pro Ser Arg Pro Asp
                245                 250                 255

aac gca acc gtg acg gtt cct gtc aac agc att tgc gag gcc gtg gcc     816
Asn Ala Thr Val Thr Val Pro Val Asn Ser Ile Cys Glu Ala Val Ala
            260                 265                 270

acc cta gac tgc acc gac gag ttt cga gaa gaa att caa agg ggc acg     864
Thr Leu Asp Cys Thr Asp Glu Phe Arg Glu Glu Ile Gln Arg Gly Thr
```

275 280 285

gcc atc ata aac tcc cag ggg cta ttg taa 894
Ala Ile Ile Asn Ser Gln Gly Leu Leu
290 295

<210> 153
<211> 297
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 153
Met Pro Lys Gln Pro Arg Ser Arg Leu Ala Ser Arg Ala Pro Tyr Ala
1 5 10 15
Pro Ser Val Arg Arg Pro Asp Gly Pro Gln Ser Thr Arg Pro Ala Ser
20 25 30
Arg His Gly Ser Cys Lys Ser Glu Ile Met Gln Trp Lys Lys Leu Val
35 40 45
Ser Asp Thr Gln Phe Phe Ser Ala Leu Thr Arg Arg His Glu Leu Gly
50 55 60
Val Asp Phe Leu Arg Glu Met Gly Thr Pro Ile Cys Thr Ser Lys Ser
65 70 75 80
Val Met Leu Pro Leu Asn Leu Lys Thr Ile Ala Pro Gly Arg Cys Val
85 90 95
Ser Leu Ser Ser Phe Gly His Ser Ser Asn Met Gly Phe Asn Cys Ser
100 105 110
Ser Cys Thr Pro Thr Asp Arg Ser Ala Val Ser Leu Asp Ala Asn Ala
115 120 125
Leu Gly Glu Asp Ser Ala Arg Lys Asn Ser Glu Leu Cys Ser Val Ala
130 135 140
Leu Thr Phe Tyr His His Ala Glu Lys Val Val Gln His Lys Gly Phe
145 150 155 160
Tyr Leu Ser Leu Leu Ser His Ser Met Glu Val Val Arg Lys Ser Phe
165 170 175
Thr Gln Pro Gly Leu Leu Tyr Ala His Leu Val Leu Lys Thr Phe Gly
180 185 190
His Asp Pro Leu Pro Ile Phe Thr Val Asp Ala Asp Glu Arg Leu Ala
195 200 205
Leu Trp Ala Val Phe His Thr Arg Asp Leu His Leu Gly Glu Thr Ser
210 215 220
Leu Arg Leu Ile Met Asp Asn Leu Pro Asn Tyr Asp Ile Thr Val Asp
225 230 235 240
Cys Ile Lys Gln Thr Tyr Ile Met Lys Phe Thr Pro Ser Arg Pro Asp
245 250 255
Asn Ala Thr Val Thr Val Pro Val Asn Ser Ile Cys Glu Ala Val Ala
260 265 270
Thr Leu Asp Cys Thr Asp Glu Phe Arg Glu Glu Ile Gln Arg Gly Thr
275 280 285
Ala Ile Ile Asn Ser Gln Gly Leu Leu
290 295

<210> 154
<211> 525
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS

<222> (1)..(525)

<400> 154
atg ttc ccg cat aag cgg tta gtg gac ttc ggt cgc cac cta gag gca 48
Met Phe Pro His Lys Arg Leu Val Asp Phe Gly Arg His Leu Glu Ala
1 5 10 15

gac gac aga gag gcc gtg ctg tgg ctt ttt gac cgg ccg gcg tcc gac 96
Asp Asp Arg Glu Ala Val Leu Trp Leu Phe Asp Arg Pro Ala Ser Asp
20 25 30

gat acg ccc gag ggg ttt gca aac ggg ctg tgc ccc tca act gga gaa 144
Asp Thr Pro Glu Gly Phe Ala Asn Gly Leu Cys Pro Ser Thr Gly Glu
35 40 45

ccc ggt att ccc ctc ccg gtg ttg ctg gag gcc gtg ttt ctc gtt ggg 192
Pro Gly Ile Pro Leu Pro Val Leu Leu Glu Ala Val Phe Leu Val Gly
50 55 60

cga ttg gac ctg gtc tcc acc ttt ttt tta cta gac gtg gga ttt att 240
Arg Leu Asp Leu Val Ser Thr Phe Phe Leu Leu Asp Val Gly Phe Ile
65 70 75 80

atc gag agg ctc cgg tcc agc ccc agt tac ttt agt cca tac aaa cac 288
Ile Glu Arg Leu Arg Ser Ser Pro Ser Tyr Phe Ser Pro Tyr Lys His
85 90 95

ctg atg ctc tcc att gac cgc cag ctc tca gaa agg gac gtg aaa aat 336
Leu Met Leu Ser Ile Asp Arg Gln Leu Ser Glu Arg Asp Val Lys Asn
100 105 110

tta gtt ttt cta acg ggc gac cag ctt ggt cgc agg cgc aac cag tca 384
Leu Val Phe Leu Thr Gly Asp Gln Leu Gly Arg Arg Arg Asn Gln Ser
115 120 125

ccc acc ttt ttt cgg tgg ctc tcg caa atg gaa aag gcg gcc ctg gtc 432
Pro Thr Phe Phe Arg Trp Leu Ser Gln Met Glu Lys Ala Ala Leu Val
130 135 140

agc ccc tca aac tac atg gtt tta agt gac ctg ctg cag gcc gtc tcc 480
Ser Pro Ser Asn Tyr Met Val Leu Ser Asp Leu Leu Gln Ala Val Ser
145 150 155 160

agg cgc gac gtg gcc aag gtc gtg gcc gcc aac gca ccc ggt taa 525
Arg Arg Asp Val Ala Lys Val Val Ala Ala Asn Ala Pro Gly
165 170 175

<210> 155
<211> 174
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 155
Met Phe Pro His Lys Arg Leu Val Asp Phe Gly Arg His Leu Glu Ala
1 5 10 15
Asp Asp Arg Glu Ala Val Leu Trp Leu Phe Asp Arg Pro Ala Ser Asp
20 25 30
Asp Thr Pro Glu Gly Phe Ala Asn Gly Leu Cys Pro Ser Thr Gly Glu
35 40 45
Pro Gly Ile Pro Leu Pro Val Leu Leu Glu Ala Val Phe Leu Val Gly

```
        50                  55                  60
Arg Leu Asp Leu Val Ser Thr Phe Phe Leu Leu Asp Val Gly Phe Ile
65                  70                  75                  80
Ile Glu Arg Leu Arg Ser Ser Pro Ser Tyr Phe Ser Pro Tyr Lys His
                85                  90                  95
Leu Met Leu Ser Ile Asp Arg Gln Leu Ser Glu Arg Asp Val Lys Asn
            100                 105                 110
Leu Val Phe Leu Thr Gly Asp Gln Leu Gly Arg Arg Arg Asn Gln Ser
        115                 120                 125
Pro Thr Phe Phe Arg Trp Leu Ser Gln Met Glu Lys Ala Ala Leu Val
    130                 135                 140
Ser Pro Ser Asn Tyr Met Val Leu Ser Asp Leu Leu Gln Ala Val Ser
145                 150                 155                 160
Arg Arg Asp Val Ala Lys Val Val Ala Ala Asn Ala Pro Gly
                165                 170
```

```
<210> 156
<211> 765
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(765)

<400> 156
atg gct tct gtt ggc cca gtg ccc acc ggg acc att gac ccc gtc ctg      48
Met Ala Ser Val Gly Pro Val Pro Thr Gly Thr Ile Asp Pro Val Leu
1               5                   10                  15

tac cag gac cgg gcc ctc tct aac ctc ctg gcc cac gaa gcg tcc ttt      96
Tyr Gln Asp Arg Ala Leu Ser Asn Leu Leu Ala His Glu Ala Ser Phe
            20                  25                  30

gta aca tcc acg gcc tgc tat ggg acg ata caa acg gag gtg acg gtg      144
Val Thr Ser Thr Ala Cys Tyr Gly Thr Ile Gln Thr Glu Val Thr Val
        35                  40                  45

ggc atg cga gtt atc ttg ggg acg tgg atg cgg tct gtt gcc agg gcc      192
Gly Met Arg Val Ile Leu Gly Thr Trp Met Arg Ser Val Ala Arg Ala
    50                  55                  60

cac cag gcg gat gct agc gtt ttc cca ctg gcc gtt agt att ctg gat      240
His Gln Ala Asp Ala Ser Val Phe Pro Leu Ala Val Ser Ile Leu Asp
65                  70                  75                  80

cgc tat cta gag tgt cgc agc atc ccc cgt cgg cgg ttt cag agg ctg      288
Arg Tyr Leu Glu Cys Arg Ser Ile Pro Arg Arg Arg Phe Gln Arg Leu
                85                  90                  95

ggg gcc gcc tgt cta ttt tta gcc ggc aag att aga gac ctg aat cct      336
Gly Ala Ala Cys Leu Phe Leu Ala Gly Lys Ile Arg Asp Leu Asn Pro
            100                 105                 110

ttc aag gcc gcc ttt tta tgc ttt tgc gcg gcg gag gat ttt tcc gtg      384
Phe Lys Ala Ala Phe Leu Cys Phe Cys Ala Ala Glu Asp Phe Ser Val
        115                 120                 125

gcg gat tta ctg aaa cag gaa aaa tca atc ctt aag gct tta agg tgg      432
```

Ala Asp Leu Leu Lys Gln Glu Lys Ser Ile Leu Lys Ala Leu Arg Trp
130 135 140

aaa cta gag gcc gtc tta cca acg gac gcc atc ggg ccc acc ctg ttt 480
Lys Leu Glu Ala Val Leu Pro Thr Asp Ala Ile Gly Pro Thr Leu Phe
145 150 155 160

aaa agc ggc ttt aca aaa gag cag ctc ttt gcc ctt cac agc cag gtg 528
Lys Ser Gly Phe Thr Lys Glu Gln Leu Phe Ala Leu His Ser Gln Val
165 170 175

gtg gag tct gtt cac aaa gcc atc gtg aac ccg gcc acc ggc ggt ctg 576
Val Glu Ser Val His Lys Ala Ile Val Asn Pro Ala Thr Gly Gly Leu
180 185 190

ccc ccg tcc ctg gtg gcg gcc gcc tgc gcg ctg ttt agc ctc ggt gcc 624
Pro Pro Ser Leu Val Ala Ala Ala Cys Ala Leu Phe Ser Leu Gly Ala
195 200 205

gct gcg ccg cct ccg gcc aga ttg gcg gag gcc gtc ggc gtt tcg gcc 672
Ala Ala Pro Pro Pro Ala Arg Leu Ala Glu Ala Val Gly Val Ser Ala
210 215 220

gca acc ctc gcg gcc gcc gcc gag tcg gtt gcc acc acc ttg cgg gaa 720
Ala Thr Leu Ala Ala Ala Ala Glu Ser Val Ala Thr Thr Leu Arg Glu
225 230 235 240

ttt gac gaa gac cac att tta agt aac gcc cgc ggt tcg tcg tga 765
Phe Asp Glu Asp His Ile Leu Ser Asn Ala Arg Gly Ser Ser
245 250 255

<210> 157
<211> 254
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 157
Met Ala Ser Val Gly Pro Val Pro Thr Gly Thr Ile Asp Pro Val Leu
1 5 10 15
Tyr Gln Asp Arg Ala Leu Ser Asn Leu Leu Ala His Glu Ala Ser Phe
20 25 30
Val Thr Ser Thr Ala Cys Tyr Gly Thr Ile Gln Thr Glu Val Thr Val
35 40 45
Gly Met Arg Val Ile Leu Gly Thr Trp Met Arg Ser Val Ala Arg Ala
50 55 60
His Gln Ala Asp Ala Ser Val Phe Pro Leu Ala Val Ser Ile Leu Asp
65 70 75 80
Arg Tyr Leu Glu Cys Arg Ser Ile Pro Arg Arg Arg Phe Gln Arg Leu
85 90 95
Gly Ala Ala Cys Leu Phe Leu Ala Gly Lys Ile Arg Asp Leu Asn Pro
100 105 110
Phe Lys Ala Ala Phe Leu Cys Phe Cys Ala Ala Glu Asp Phe Ser Val
115 120 125
Ala Asp Leu Leu Lys Gln Glu Lys Ser Ile Leu Lys Ala Leu Arg Trp
130 135 140
Lys Leu Glu Ala Val Leu Pro Thr Asp Ala Ile Gly Pro Thr Leu Phe
145 150 155 160
Lys Ser Gly Phe Thr Lys Glu Gln Leu Phe Ala Leu His Ser Gln Val
165 170 175
Val Glu Ser Val His Lys Ala Ile Val Asn Pro Ala Thr Gly Gly Leu

180 185 190
Pro Pro Ser Leu Val Ala Ala Ala Cys Ala Leu Phe Ser Leu Gly Ala
195 200 205
Ala Ala Pro Pro Pro Ala Arg Leu Ala Glu Ala Val Gly Val Ser Ala
210 215 220
Ala Thr Leu Ala Ala Ala Ala Glu Ser Val Ala Thr Thr Leu Arg Glu
225 230 235 240
Phe Asp Glu Asp His Ile Leu Ser Asn Ala Arg Gly Ser Ser
245 250

<210> 158
<211> 1347
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1347)

<400> 158
atg tgg ggc agc cgg caa cat cgt tcc ggc ata gta tcg gga cac ggg 48
Met Trp Gly Ser Arg Gln His Arg Ser Gly Ile Val Ser Gly His Gly
1 5 10 15

ttg cga agt tcc tgc agg gga cat tgc gga aga cgc ggc ggc act aga 96
Leu Arg Ser Ser Cys Arg Gly His Cys Gly Arg Arg Gly Gly Thr Arg
20 25 30

gaa cag gcc ggg agg cgc ggc agg ggg cgt ggc acg gcc gcg cca gcc 144
Glu Gln Ala Gly Arg Arg Gly Arg Gly Arg Gly Thr Ala Ala Pro Ala
35 40 45

gcc gca ccc gcg ccc ccc gct cca aca aca tca ggg ccg caa gtg cgg 192
Ala Ala Pro Ala Pro Pro Ala Pro Thr Thr Ser Gly Pro Gln Val Arg
50 55 60

gcg gtt gca gaa caa ggg cac gga tcg gac acc gaa acg gcc acc gaa 240
Ala Val Ala Glu Gln Gly His Gly Ser Asp Thr Glu Thr Ala Thr Glu
65 70 75 80

tcc agg cac ggg tca tcc caa ggg tcg cca tca ggg tca ggg tct gag 288
Ser Arg His Gly Ser Ser Gln Gly Ser Pro Ser Gly Ser Gly Ser Glu
85 90 95

tcc gtc atc gtg ttg gga tct cca acg ccg tcg cca agc ggg tct gcg 336
Ser Val Ile Val Leu Gly Ser Pro Thr Pro Ser Pro Ser Gly Ser Ala
100 105 110

ccg gtg tta gct tca ggc ctg tca cct caa aac acg tca ggg tcg tcc 384
Pro Val Leu Ala Ser Gly Leu Ser Pro Gln Asn Thr Ser Gly Ser Ser
115 120 125

ccg gca tcc ccc gca tcg cat tcg ccc cca cca tca ccc ccg tca cac 432
Pro Ala Ser Pro Ala Ser His Ser Pro Pro Pro Ser Pro Pro Ser His
130 135 140

ccc ggc ccc cac tca cct gca ccg ccg tcg tcg cat aat cca tcg ccc 480
Pro Gly Pro His Ser Pro Ala Pro Pro Ser Ser His Asn Pro Ser Pro
145 150 155 160

```
aat cag cag cct tcg tcg ttt ttg caa cca tcg cac cac gac tca cct      528
Asn Gln Gln Pro Ser Ser Phe Leu Gln Pro Ser His His Asp Ser Pro
                165                 170                 175

gaa ccg cct gag cct ccg acg tcc ttg ccc cct ccc gac tca cct gga      576
Glu Pro Pro Glu Pro Pro Thr Ser Leu Pro Pro Pro Asp Ser Pro Gly
            180                 185                 190

ccg cca caa tcg cca acg ccg acg tcc tcg ccc cct ccg caa tcc ccg      624
Pro Pro Gln Ser Pro Thr Pro Thr Ser Ser Pro Pro Pro Gln Ser Pro
        195                 200                 205

ccc gac tca ccg gga cca ccc caa tcg cca acg cct caa cag gcc cca      672
Pro Asp Ser Pro Gly Pro Pro Gln Ser Pro Thr Pro Gln Gln Ala Pro
    210                 215                 220

tcc cct aac acg caa cag gcg gtg agc cac aca gac cac ccc acc ggg      720
Ser Pro Asn Thr Gln Gln Ala Val Ser His Thr Asp His Pro Thr Gly
225                 230                 235                 240

ccg tct cgg cca gga ccc ccg ttt cct ggc cat aca tcc cac tcg tac      768
Pro Ser Arg Pro Gly Pro Pro Phe Pro Gly His Thr Ser His Ser Tyr
                245                 250                 255

acc gta ggg ggt tgg ggt cca ccc acg cgt gcc gga gga gtt cct tgt      816
Thr Val Gly Gly Trp Gly Pro Pro Thr Arg Ala Gly Gly Val Pro Cys
            260                 265                 270

tta cgt ctc aga tgc acg agt cac aac tcc cac gag gac gaa gct cct      864
Leu Arg Leu Arg Cys Thr Ser His Asn Ser His Glu Asp Glu Ala Pro
        275                 280                 285

gaa aga caa cag gag caa gag ggg gag gag cgg cag cag caa ccg gcg      912
Glu Arg Gln Gln Glu Gln Glu Gly Glu Glu Arg Gln Gln Gln Pro Ala
    290                 295                 300

cgc ccc ccg cgc ccc ccg cgc ccg ccg cga tac ccc att ccc ata ccg      960
Arg Pro Pro Arg Pro Pro Arg Pro Pro Arg Tyr Pro Ile Pro Ile Pro
305                 310                 315                 320

tat ccc tca tca gag gaa gag gta ccc aga aaa tac cgc cca cag aga      1008
Tyr Pro Ser Ser Glu Glu Glu Val Pro Arg Lys Tyr Arg Pro Gln Arg
                325                 330                 335

aga ttt tat cgt cag gta ttg gga ccc agg atc gac ccc cct cgg ccg      1056
Arg Phe Tyr Arg Gln Val Leu Gly Pro Arg Ile Asp Pro Pro Arg Pro
            340                 345                 350

ggc ccg tgg tgt cac gga gtc atc ttt tgt aac tca gac ccc tat tca      1104
Gly Pro Trp Cys His Gly Val Ile Phe Cys Asn Ser Asp Pro Tyr Ser
        355                 360                 365

ctg tat cgc ctg gca agg tgc cta cag ttt ccc ggc att aga gca tcc      1152
Leu Tyr Arg Leu Ala Arg Cys Leu Gln Phe Pro Gly Ile Arg Ala Ser
    370                 375                 380

agc gtc cgc gtt ttg ccg gac gca cct ggc agt ccc gtc atc ccg gct      1200
Ser Val Arg Val Leu Pro Asp Ala Pro Gly Ser Pro Val Ile Pro Ala
385                 390                 395                 400
```

ttt tgc att acg gtg ttc tgt caa agc cgc gga acc gcc aag gcc gtc 1248
Phe Cys Ile Thr Val Phe Cys Gln Ser Arg Gly Thr Ala Lys Ala Val
405 410 415

aaa aag gcc cgt cgc cgg tgg gaa cga cat cac cca tcc gcc ccc cac 1296
Lys Lys Ala Arg Arg Arg Trp Glu Arg His His Pro Ser Ala Pro His
420 425 430

ttc cag gcc agc ata gtg cgc atg gac aga gga ctg cca att cag cac 1344
Phe Gln Ala Ser Ile Val Arg Met Asp Arg Gly Leu Pro Ile Gln His
435 440 445

taa 1347

<210> 159
<211> 448
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 159
Met Trp Gly Ser Arg Gln His Arg Ser Gly Ile Val Ser Gly His Gly
1 5 10 15
Leu Arg Ser Ser Cys Arg Gly His Cys Gly Arg Arg Gly Gly Thr Arg
20 25 30
Glu Gln Ala Gly Arg Arg Gly Arg Gly Arg Gly Thr Ala Ala Pro Ala
35 40 45
Ala Ala Pro Ala Pro Pro Ala Pro Thr Thr Ser Gly Pro Gln Val Arg
50 55 60
Ala Val Ala Glu Gln Gly His Gly Ser Asp Thr Glu Thr Ala Thr Glu
65 70 75 80
Ser Arg His Gly Ser Ser Gln Gly Ser Pro Ser Gly Ser Gly Ser Glu
85 90 95
Ser Val Ile Val Leu Gly Ser Pro Thr Pro Ser Pro Ser Gly Ser Ala
100 105 110
Pro Val Leu Ala Ser Gly Leu Ser Pro Gln Asn Thr Ser Gly Ser Ser
115 120 125
Pro Ala Ser Pro Ala Ser His Ser Pro Pro Pro Ser Pro Pro Ser His
130 135 140
Pro Gly Pro His Ser Pro Ala Pro Pro Ser Ser His Asn Pro Ser Pro
145 150 155 160
Asn Gln Gln Pro Ser Ser Phe Leu Gln Pro Ser His His Asp Ser Pro
165 170 175
Glu Pro Pro Glu Pro Pro Thr Ser Leu Pro Pro Pro Asp Ser Pro Gly
180 185 190
Pro Pro Gln Ser Pro Thr Pro Thr Ser Ser Pro Pro Pro Gln Ser Pro
195 200 205
Pro Asp Ser Pro Gly Pro Pro Gln Ser Pro Thr Pro Gln Gln Ala Pro
210 215 220
Ser Pro Asn Thr Gln Gln Ala Val Ser His Thr Asp His Pro Thr Gly
225 230 235 240
Pro Ser Arg Pro Gly Pro Pro Phe Pro Gly His Thr Ser His Ser Tyr
245 250 255
Thr Val Gly Gly Trp Gly Pro Pro Thr Arg Ala Gly Gly Val Pro Cys
260 265 270
Leu Arg Leu Arg Cys Thr Ser His Asn Ser His Glu Asp Glu Ala Pro
275 280 285
Glu Arg Gln Gln Glu Gln Glu Gly Glu Glu Arg Gln Gln Gln Pro Ala
290 295 300
Arg Pro Pro Arg Pro Pro Arg Pro Pro Arg Tyr Pro Ile Pro Ile Pro

```
305                 310                 315                 320
Tyr Pro Ser Ser Glu Glu Glu Val Pro Arg Lys Tyr Arg Pro Gln Arg
                325                 330                 335
Arg Phe Tyr Arg Gln Val Leu Gly Pro Arg Ile Asp Pro Pro Arg Pro
            340                 345                 350
Gly Pro Trp Cys His Gly Val Ile Phe Cys Asn Ser Asp Pro Tyr Ser
        355                 360                 365
Leu Tyr Arg Leu Ala Arg Cys Leu Gln Phe Pro Gly Ile Arg Ala Ser
    370                 375                 380
Ser Val Arg Val Leu Pro Asp Ala Pro Gly Ser Pro Val Ile Pro Ala
385                 390                 395                 400
Phe Cys Ile Thr Val Phe Cys Gln Ser Arg Gly Thr Ala Lys Ala Val
                405                 410                 415
Lys Lys Ala Arg Arg Arg Trp Glu Arg His His Pro Ser Ala Pro His
            420                 425                 430
Phe Gln Ala Ser Ile Val Arg Met Asp Arg Gly Leu Pro Ile Gln His
        435                 440                 445
```

```
<210> 160
<211> 762
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(762)

<400> 160
atg tcg gga gga att aca tta acg ctg ctg ctg gcg acg ctt gcc acg      48
Met Ser Gly Gly Ile Thr Leu Thr Leu Leu Leu Ala Thr Leu Ala Thr
1               5                   10                  15

gtt cgg tgc gct ctt caa acg cac tat gcg gcg gtc ccc gtg cac tct      96
Val Arg Cys Ala Leu Gln Thr His Tyr Ala Ala Val Pro Val His Ser
            20                  25                  30

acc gcg tcc ctg ggg tgc gtg tta acg aca ccc cac gac gtt ctt atc     144
Thr Ala Ser Leu Gly Cys Val Leu Thr Thr Pro His Asp Val Leu Ile
        35                  40                  45

gtt acc tgg caa aaa cag gaa tcg cct agt ccc gtt aac gtg gcc aca     192
Val Thr Trp Gln Lys Gln Glu Ser Pro Ser Pro Val Asn Val Ala Thr
    50                  55                  60

tat agt tcc gaa gcg ggc acg gtg gtt cag ccc ccg ttc gcc ggt agg     240
Tyr Ser Ser Glu Ala Gly Thr Val Val Gln Pro Pro Phe Ala Gly Arg
65                  70                  75                  80

gtt gac att ccc gaa cac aag ttg acc aga acg acc ctg aag ttt ttt     288
Val Asp Ile Pro Glu His Lys Leu Thr Arg Thr Thr Leu Lys Phe Phe
                85                  90                  95

aat gcc acc ctg gag gac gag ggg tgc tac ctg tgt atc ttt aac gcg     336
Asn Ala Thr Leu Glu Asp Glu Gly Cys Tyr Leu Cys Ile Phe Asn Ala
            100                 105                 110

ttt gga gtg gga aag ctg tcg gga acc gcc tgc ttg acg gtt tac gtc     384
Phe Gly Val Gly Lys Leu Ser Gly Thr Ala Cys Leu Thr Val Tyr Val
        115                 120                 125
```

```
ccc ctg tcc atg tcc gtc acg ttt tac ccc ccg att aac ccg acg cag     432
Pro Leu Ser Met Ser Val Thr Phe Tyr Pro Pro Ile Asn Pro Thr Gln
    130                 135                 140

ctc gtc tgt cgg gcc gag gcc agt ccc gca ccg tcg gtc aac tgg acc     480
Leu Val Cys Arg Ala Glu Ala Ser Pro Ala Pro Ser Val Asn Trp Thr
145                 150                 155                 160

ggc gtg ccg ccc gag ctg tgc agc gaa cct gaa gtg ttt ccc cgg ccc     528
Gly Val Pro Pro Glu Leu Cys Ser Glu Pro Glu Val Phe Pro Arg Pro
                165                 170                 175

aac gga aca acc ctg gtt gtg ggt cgc tgc aac gta acg tcg gtg gac     576
Asn Gly Thr Thr Leu Val Val Gly Arg Cys Asn Val Thr Ser Val Asp
            180                 185                 190

ccc gaa gac ctt gag aac gcc acg tgc ctg gtc act cat ata ggc ggt     624
Pro Glu Asp Leu Glu Asn Ala Thr Cys Leu Val Thr His Ile Gly Gly
        195                 200                 205

ttg gcc gcg gcg cgg ccc ctg gac ccc gtg ttt tcg gat ccc ctg gaa     672
Leu Ala Ala Ala Arg Pro Leu Asp Pro Val Phe Ser Asp Pro Leu Glu
    210                 215                 220

ggg acg agc cac tac gtg gtg ggt gtg gtg gca gcg gcc gcc gtt tta     720
Gly Thr Ser His Tyr Val Val Gly Val Val Ala Ala Ala Ala Val Leu
225                 230                 235                 240

ggc att ttt tta acg ggt gtt ttt ttg tat agg tct atg tga             762
Gly Ile Phe Leu Thr Gly Val Phe Leu Tyr Arg Ser Met
                245                 250


<210> 161
<211> 253
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 161
Met Ser Gly Gly Ile Thr Leu Thr Leu Leu Leu Ala Thr Leu Ala Thr
1               5                   10                  15
Val Arg Cys Ala Leu Gln Thr His Tyr Ala Ala Val Pro Val His Ser
            20                  25                  30
Thr Ala Ser Leu Gly Cys Val Leu Thr Thr Pro His Asp Val Leu Ile
        35                  40                  45
Val Thr Trp Gln Lys Gln Glu Ser Pro Ser Pro Val Asn Val Ala Thr
    50                  55                  60
Tyr Ser Ser Glu Ala Gly Thr Val Val Gln Pro Pro Phe Ala Gly Arg
65                  70                  75                  80
Val Asp Ile Pro Glu His Lys Leu Thr Arg Thr Thr Leu Lys Phe Phe
                85                  90                  95
Asn Ala Thr Leu Glu Asp Glu Gly Cys Tyr Leu Cys Ile Phe Asn Ala
            100                 105                 110
Phe Gly Val Gly Lys Leu Ser Gly Thr Ala Cys Leu Thr Val Tyr Val
        115                 120                 125
Pro Leu Ser Met Ser Val Thr Phe Tyr Pro Pro Ile Asn Pro Thr Gln
    130                 135                 140
Leu Val Cys Arg Ala Glu Ala Ser Pro Ala Pro Ser Val Asn Trp Thr
145                 150                 155                 160
Gly Val Pro Pro Glu Leu Cys Ser Glu Pro Glu Val Phe Pro Arg Pro
```

165 170 175
Asn Gly Thr Thr Leu Val Val Gly Arg Cys Asn Val Thr Ser Val Asp
180 185 190
Pro Glu Asp Leu Glu Asn Ala Thr Cys Leu Val Thr His Ile Gly Gly
195 200 205
Leu Ala Ala Ala Arg Pro Leu Asp Pro Val Phe Ser Asp Pro Leu Glu
210 215 220
Gly Thr Ser His Tyr Val Val Gly Val Val Ala Ala Ala Ala Val Leu
225 230 235 240
Gly Ile Phe Leu Thr Gly Val Phe Leu Tyr Arg Ser Met
245 250

<210> 162
<211> 1029
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(1029)

<400> 162
atg gac gcc ttg aac aat aac ctt aac ctg ctg atg gat ttt ctg tct 48
Met Asp Ala Leu Asn Asn Asn Leu Asn Leu Leu Met Asp Phe Leu Ser
1 5 10 15

aac tat tcg aat agc tac agt agt tat gac gac aat atg tct tac acc 96
Asn Tyr Ser Asn Ser Tyr Ser Ser Tyr Asp Asp Asn Met Ser Tyr Thr
20 25 30

tta gac acg gaa tcc acg ctg tgt cgg ctg acg gtg gtt ttc cca cct 144
Leu Asp Thr Glu Ser Thr Leu Cys Arg Leu Thr Val Val Phe Pro Pro
35 40 45

acc gtt tat gcg att ata tgc ttt ttt att ttt tgc att acg ctg ttt 192
Thr Val Tyr Ala Ile Ile Cys Phe Phe Ile Phe Cys Ile Thr Leu Phe
50 55 60

ggg aac gcg ttg gtg cta tat att ttt ttt aaa ttt aaa gcg ctc gcc 240
Gly Asn Ala Leu Val Leu Tyr Ile Phe Phe Lys Phe Lys Ala Leu Ala
65 70 75 80

aac tct gtg gat gta ctg atg gct ggg ttg tgt tgt aac tcc ctg ttt 288
Asn Ser Val Asp Val Leu Met Ala Gly Leu Cys Cys Asn Ser Leu Phe
85 90 95

ctg tgc gcg tcg ttt ttg ttc agc tgg ctg ctg tac gtc gcg cca cag 336
Leu Cys Ala Ser Phe Leu Phe Ser Trp Leu Leu Tyr Val Ala Pro Gln
100 105 110

atg ctc acg tcc gcg acg tgc aag gtg gaa atc ttt ttc ttt tac ctg 384
Met Leu Thr Ser Ala Thr Cys Lys Val Glu Ile Phe Phe Phe Tyr Leu
115 120 125

tac acg tac ttt ggc gtg tac att gtg gtg tgt atc agc ctt atc agg 432
Tyr Thr Tyr Phe Gly Val Tyr Ile Val Val Cys Ile Ser Leu Ile Arg
130 135 140

tgc ctg tta gtt gtg ttt tcc cgc cgc ccg tgg gtc aag cac ggg gcc 480

Cys Leu Leu Val Val Phe Ser Arg Arg Pro Trp Val Lys His Gly Ala
145 150 155 160

tcc ggc ttt ctc tgc gtg tgt gtg tct tta atc gtg gcg ctg gcg ctg 528
Ser Gly Phe Leu Cys Val Cys Val Ser Leu Ile Val Ala Leu Ala Leu
165 170 175

tct gcc aac gcg agc ctc tat agg acg gcc ctg cgt cac cca gag acc 576
Ser Ala Asn Ala Ser Leu Tyr Arg Thr Ala Leu Arg His Pro Glu Thr
180 185 190

agc gag tgg ata tgc tac gaa gat gcc ggg gaa gat acc gtc aac tgg 624
Ser Glu Trp Ile Cys Tyr Glu Asp Ala Gly Glu Asp Thr Val Asn Trp
195 200 205

aag ctg aga atc aga acc acc agc gcg atc tgc ggg ttt ttg gtt ccg 672
Lys Leu Arg Ile Arg Thr Thr Ser Ala Ile Cys Gly Phe Leu Val Pro
210 215 220

ttt ggg ctg atg gtg ctc ttt tac gga ctt acg tgg tgt atg gtt aaa 720
Phe Gly Leu Met Val Leu Phe Tyr Gly Leu Thr Trp Cys Met Val Lys
225 230 235 240

agc acg aag ctg gcc aga aag gga gcc gtt agg ggt gta att gtg acg 768
Ser Thr Lys Leu Ala Arg Lys Gly Ala Val Arg Gly Val Ile Val Thr
245 250 255

gtg gtg gtg ctg ttt tta att ttt tgc ctg ccc tat cac ctg tgc aac 816
Val Val Val Leu Phe Leu Ile Phe Cys Leu Pro Tyr His Leu Cys Asn
260 265 270

ttt ttt gac acc ctg ttg agg acc ggt ttt ctg gcc gaa acg tgc tac 864
Phe Phe Asp Thr Leu Leu Arg Thr Gly Phe Leu Ala Glu Thr Cys Tyr
275 280 285

ctc agg gac gtg atc agc gtg gcc atg cac ata tgc tcc ctg cta cag 912
Leu Arg Asp Val Ile Ser Val Ala Met His Ile Cys Ser Leu Leu Gln
290 295 300

agc atg tat agc gcg ttc gtg cca gtc gtg tat tct ggt ctt ggg tct 960
Ser Met Tyr Ser Ala Phe Val Pro Val Val Tyr Ser Gly Leu Gly Ser
305 310 315 320

ctg ttt agg aga agg gtt agg gat acc tgg tcc gtg ttt agg tgt ttt 1008
Leu Phe Arg Arg Arg Val Arg Asp Thr Trp Ser Val Phe Arg Cys Phe
325 330 335

tcc act tca ggt agt tta tga 1029
Ser Thr Ser Gly Ser Leu
340

<210> 163
<211> 342
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 163
Met Asp Ala Leu Asn Asn Asn Leu Asn Leu Leu Met Asp Phe Leu Ser
1 5 10 15
Asn Tyr Ser Asn Ser Tyr Ser Ser Tyr Asp Asp Asn Met Ser Tyr Thr

20 25 30
Leu Asp Thr Glu Ser Thr Leu Cys Arg Leu Thr Val Val Phe Pro Pro
35 40 45
Thr Val Tyr Ala Ile Ile Cys Phe Phe Ile Phe Cys Ile Thr Leu Phe
50 55 60
Gly Asn Ala Leu Val Leu Tyr Ile Phe Phe Lys Phe Lys Ala Leu Ala
65 70 75 80
Asn Ser Val Asp Val Leu Met Ala Gly Leu Cys Cys Asn Ser Leu Phe
85 90 95
Leu Cys Ala Ser Phe Leu Phe Ser Trp Leu Leu Tyr Val Ala Pro Gln
100 105 110
Met Leu Thr Ser Ala Thr Cys Lys Val Glu Ile Phe Phe Phe Tyr Leu
115 120 125
Tyr Thr Tyr Phe Gly Val Tyr Ile Val Val Cys Ile Ser Leu Ile Arg
130 135 140
Cys Leu Leu Val Val Phe Ser Arg Arg Pro Trp Val Lys His Gly Ala
145 150 155 160
Ser Gly Phe Leu Cys Val Cys Val Ser Leu Ile Val Ala Leu Ala Leu
165 170 175
Ser Ala Asn Ala Ser Leu Tyr Arg Thr Ala Leu Arg His Pro Glu Thr
180 185 190
Ser Glu Trp Ile Cys Tyr Glu Asp Ala Gly Glu Asp Thr Val Asn Trp
195 200 205
Lys Leu Arg Ile Arg Thr Thr Ser Ala Ile Cys Gly Phe Leu Val Pro
210 215 220
Phe Gly Leu Met Val Leu Phe Tyr Gly Leu Thr Trp Cys Met Val Lys
225 230 235 240
Ser Thr Lys Leu Ala Arg Lys Gly Ala Val Arg Gly Val Ile Val Thr
245 250 255
Val Val Val Leu Phe Leu Ile Phe Cys Leu Pro Tyr His Leu Cys Asn
260 265 270
Phe Phe Asp Thr Leu Leu Arg Thr Gly Phe Leu Ala Glu Thr Cys Tyr
275 280 285
Leu Arg Asp Val Ile Ser Val Ala Met His Ile Cys Ser Leu Leu Gln
290 295 300
Ser Met Tyr Ser Ala Phe Val Pro Val Val Tyr Ser Gly Leu Gly Ser
305 310 315 320
Leu Phe Arg Arg Arg Val Arg Asp Thr Trp Ser Val Phe Arg Cys Phe
325 330 335
Ser Thr Ser Gly Ser Leu
340

<210> 164
<211> 3897
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<220>
<221> CDS
<222> (1)..(3897)

<400> 164
atg gcc caa cgc acg aac cct cgc tgg gcc gct gca gcc ctt agt cca 48
Met Ala Gln Arg Thr Asn Pro Arg Trp Ala Ala Ala Ala Leu Ser Pro
1 5 10 15

gaa gag gag gct ttt att cat gac aat tcg gac gca gaa agt gtt ttg 96
Glu Glu Glu Ala Phe Ile His Asp Asn Ser Asp Ala Glu Ser Val Leu
20 25 30

```
gcg ctt gtg cca gaa cag tgc ttt agt gaa ttt tta ctt tgg ctt gta    144
Ala Leu Val Pro Glu Gln Cys Phe Ser Glu Phe Leu Leu Trp Leu Val
        35                  40                  45

aca cgc ccc tct gat aac ttt gac aac gac gat gac gac cca gcg ctg    192
Thr Arg Pro Ser Asp Asn Phe Asp Asn Asp Asp Asp Asp Pro Ala Leu
    50                  55                  60

ggc gtt atc tgg cat ctt ctg gcg cct ctg gtt aat tat gca cct ctg    240
Gly Val Ile Trp His Leu Leu Ala Pro Leu Val Asn Tyr Ala Pro Leu
65                  70                  75                  80

gaa act cgg tcg gcg cac ctc cag ggc cat cat act ata tcc ctg ccc    288
Glu Thr Arg Ser Ala His Leu Gln Gly His His Thr Ile Ser Leu Pro
                85                  90                  95

tat ggc cca gac ctg atg cgc caa cct acc aca aga tct agc gaa ata    336
Tyr Gly Pro Asp Leu Met Arg Gln Pro Thr Thr Arg Ser Ser Glu Ile
            100                 105                 110

gtg cag tgc ctt aga gac agc ggc ctc gat aga acg ttg cgg tta gag    384
Val Gln Cys Leu Arg Asp Ser Gly Leu Asp Arg Thr Leu Arg Leu Glu
        115                 120                 125

gtg ggc aga cat ctg agc tgc cag acg aga cgg ttt gtc gcc gat cgg    432
Val Gly Arg His Leu Ser Cys Gln Thr Arg Arg Phe Val Ala Asp Arg
    130                 135                 140

gta ccc ccg ggc acc ttg gcc gcc ctg aca ctt ggc aca cta gta gaa    480
Val Pro Pro Gly Thr Leu Ala Ala Leu Thr Leu Gly Thr Leu Val Glu
145                 150                 155                 160

tat gat gtg cgc gtg cag cgc cag ctc ccg gtg aca ttg caa tcc acc    528
Tyr Asp Val Arg Val Gln Arg Gln Leu Pro Val Thr Leu Gln Ser Thr
                165                 170                 175

gcc tgg aga ccg ttg ccc gag aga gac cca ata tgc gcc gcg gtg atg    576
Ala Trp Arg Pro Leu Pro Glu Arg Asp Pro Ile Cys Ala Ala Val Met
            180                 185                 190

ctc ccg tta caa cgg aac ata tta ccg ctg gcc gtg cag gcc tcc aac    624
Leu Pro Leu Gln Arg Asn Ile Leu Pro Leu Ala Val Gln Ala Ser Asn
        195                 200                 205

ggc aac agc tat acg gtg tcc aga tac gcc gtc atg gcc cgc agg agc    672
Gly Asn Ser Tyr Thr Val Ser Arg Tyr Ala Val Met Ala Arg Arg Ser
    210                 215                 220

tac agc tgc gtt ttc cag cgc ctc ccg tgc gaa aac gta acc cac ata    720
Tyr Ser Cys Val Phe Gln Arg Leu Pro Cys Glu Asn Val Thr His Ile
225                 230                 235                 240

gct gac tca ttt aca cac ctg cac agc gcc att cag aca ggt gca ggt    768
Ala Asp Ser Phe Thr His Leu His Ser Ala Ile Gln Thr Gly Ala Gly
                245                 250                 255

gcg ctg caa aac att ctg ttc cat gcc acg ctg ctg ccc ggg ggc gaa    816
Ala Leu Gln Asn Ile Leu Phe His Ala Thr Leu Leu Pro Gly Gly Glu
            260                 265                 270
```

atc aga tcg gcc ctg tgt gga ttt tac gcc act acg ccg tca gtg ggc 864
Ile Arg Ser Ala Leu Cys Gly Phe Tyr Ala Thr Thr Pro Ser Val Gly
275 280 285

gca ttt tct cgc gca cgc cac aga gct att aac aca aca gcg aca ctc 912
Ala Phe Ser Arg Ala Arg His Arg Ala Ile Asn Thr Thr Ala Thr Leu
290 295 300

cac tgc cag cag ctg gcg cgc acc ggc acg cct gtc ctc ggt ggc ttt 960
His Cys Gln Gln Leu Ala Arg Thr Gly Thr Pro Val Leu Gly Gly Phe
305 310 315 320

ctt aaa acc gtc cac agc gcc acc acc agc gag gcg aac gtt att acc 1008
Leu Lys Thr Val His Ser Ala Thr Thr Ser Glu Ala Asn Val Ile Thr
325 330 335

acc aca tcg ctg tta tcg tgc gtg cct caa gca tac aca ttc ctc agg 1056
Thr Thr Ser Leu Leu Ser Cys Val Pro Gln Ala Tyr Thr Phe Leu Arg
340 345 350

agg tct tta ttc agt cag cct atc atc tgt ctt ggg tct ttt gaa ccc 1104
Arg Ser Leu Phe Ser Gln Pro Ile Ile Cys Leu Gly Ser Phe Glu Pro
355 360 365

gtt gac ggc gat ggc aac cag cgc tcg ctt tac ctg ggg agc gcc gca 1152
Val Asp Gly Asp Gly Asn Gln Arg Ser Leu Tyr Leu Gly Ser Ala Ala
370 375 380

ggt att acc cgc atc acc caa acg ttg tcg ctg gct tac gag att ttg 1200
Gly Ile Thr Arg Ile Thr Gln Thr Leu Ser Leu Ala Tyr Glu Ile Leu
385 390 395 400

gaa ggg ccc cta ttt acc agc att aat cgc gcc cat gaa ccc gcc tct 1248
Glu Gly Pro Leu Phe Thr Ser Ile Asn Arg Ala His Glu Pro Ala Ser
405 410 415

gtc atc ggc cac ctg gga gcc ctg gtc tcg cgg ggc ggc ctg cgc ctc 1296
Val Ile Gly His Leu Gly Ala Leu Val Ser Arg Gly Gly Leu Arg Leu
420 425 430

ttt gtc tct cag ctt cca cca acc att ctg agc caa ctg acc gcc acg 1344
Phe Val Ser Gln Leu Pro Pro Thr Ile Leu Ser Gln Leu Thr Ala Thr
435 440 445

cca gac atc tca cgg gaa acc gtg aac gac atc cta gtt aac aag ttt 1392
Pro Asp Ile Ser Arg Glu Thr Val Asn Asp Ile Leu Val Asn Lys Phe
450 455 460

ctc aac gtg tct gcc tgc gtg gtc ttt gcc gtc ctc ccg cgc gac acg 1440
Leu Asn Val Ser Ala Cys Val Val Phe Ala Val Leu Pro Arg Asp Thr
465 470 475 480

gag ccg gaa ccg ggc ccg ttg gat gcc atc agg agg gcc gca cgc atc 1488
Glu Pro Glu Pro Gly Pro Leu Asp Ala Ile Arg Arg Ala Ala Arg Ile
485 490 495

tgc gga tgc cct ttc gcc gtc gtt ggg gaa acc tgc gaa gag ctt gga 1536
Cys Gly Cys Pro Phe Ala Val Val Gly Glu Thr Cys Glu Glu Leu Gly
500 505 510

att cag ttc gtg aac gac ctg gag ctg tgg aac ccg gga gcg tgg ccg 1584

Ile Gln Phe Val Asn Asp Leu Glu Leu Trp Asn Pro Gly Ala Trp Pro
515 520 525

atc aga cag cca aca tcg gcc gaa gtc atc gca act ttt ggg ttc gac 1632
Ile Arg Gln Pro Thr Ser Ala Glu Val Ile Ala Thr Phe Gly Phe Asp
530 535 540

gag cag ccc gtt tcc tcc aac tgg ctg gtg cgc cca gaa gaa cct gag 1680
Glu Gln Pro Val Ser Ser Asn Trp Leu Val Arg Pro Glu Glu Pro Glu
545 550 555 560

gat ggt ggc gaa cag gca ccc tcg ccg acc gac tgg ggc cta ttc cgc 1728
Asp Gly Gly Glu Gln Ala Pro Ser Pro Thr Asp Trp Gly Leu Phe Arg
565 570 575

ctg gcc tcc gtg gtc gat cag ctt ctg cga tgt ccg acg gtt ggc agc 1776
Leu Ala Ser Val Val Asp Gln Leu Leu Arg Cys Pro Thr Val Gly Ser
580 585 590

aaa gag ttt gtc acg cga cac gtg gac aga tgc tcc aac gga ctc gtg 1824
Lys Glu Phe Val Thr Arg His Val Asp Arg Cys Ser Asn Gly Leu Val
595 600 605

gct cag cag tgc gaa gtg gga ccc ctg ggc cgg ccg ctg tca gat tac 1872
Ala Gln Gln Cys Glu Val Gly Pro Leu Gly Arg Pro Leu Ser Asp Tyr
610 615 620

cac att gtc aac cac acg tcg gtg ttt acg gac aga atg gcg cgg gtg 1920
His Ile Val Asn His Thr Ser Val Phe Thr Asp Arg Met Ala Arg Val
625 630 635 640

ccc ata tat cgc ccc cag ccg atc acc agg cag gac gcg acg gaa cgc 1968
Pro Ile Tyr Arg Pro Gln Pro Ile Thr Arg Gln Asp Ala Thr Glu Arg
645 650 655

ctg gtt agc cca gaa acc tgg gtc acc cag ggc agg ggc agg aac cgg 2016
Leu Val Ser Pro Glu Thr Trp Val Thr Gln Gly Arg Gly Arg Asn Arg
660 665 670

tgg gtc gga cag tgc gtg gct tat gga gaa cag gca tac aag atg ggc 2064
Trp Val Gly Gln Cys Val Ala Tyr Gly Glu Gln Ala Tyr Lys Met Gly
675 680 685

atc aac gcg gca gtg ggc gcg aga tac gcc atc tgc gag gcg gtc acc 2112
Ile Asn Ala Ala Val Gly Ala Arg Tyr Ala Ile Cys Glu Ala Val Thr
690 695 700

aac atc atg cta gcg cac gtg cgg cgt cta agc gac atc acg ctg acg 2160
Asn Ile Met Leu Ala His Val Arg Arg Leu Ser Asp Ile Thr Leu Thr
705 710 715 720

gcg tcg gtc ggt tgg aac ccg gag gac gac cag gcc tgg ctc ctg cag 2208
Ala Ser Val Gly Trp Asn Pro Glu Asp Asp Gln Ala Trp Leu Leu Gln
725 730 735

cac aca ctg ttt gcc tgc aag gaa cta tgc agg gac ctg agc atc aac 2256
His Thr Leu Phe Ala Cys Lys Glu Leu Cys Arg Asp Leu Ser Ile Asn
740 745 750

ttc gcc atc acg tcg gcc ggc agc acc ccg tgc ctg tcg gaa gaa ctg 2304
Phe Ala Ile Thr Ser Ala Gly Ser Thr Pro Cys Leu Ser Glu Glu Leu

755 760 765

atc agc gcc acc caa cag cat cag acg gta gcc ccc gtg ccg ttt aac 2352
Ile Ser Ala Thr Gln Gln His Gln Thr Val Ala Pro Val Pro Phe Asn
770 775 780

gca gtc gtc atc acg gcc acc gcc gag gtg aag tcg tcc agg caa cgc 2400
Ala Val Val Ile Thr Ala Thr Ala Glu Val Lys Ser Ser Arg Gln Arg
785 790 795 800

gtc aca cct gac ctg aag gcc acc ggt aac ctc atc gtg ctg gtg tcc 2448
Val Thr Pro Asp Leu Lys Ala Thr Gly Asn Leu Ile Val Leu Val Ser
805 810 815

ttc ccc gtg ccg cac ctg acc cag gga tcg acc ttt gag cac ctg tgt 2496
Phe Pro Val Pro His Leu Thr Gln Gly Ser Thr Phe Glu His Leu Cys
820 825 830

ctc ttg ccg agt ccc acc cta ccg gac gtt cag gcg aca cac ctg gcc 2544
Leu Leu Pro Ser Pro Thr Leu Pro Asp Val Gln Ala Thr His Leu Ala
835 840 845

aac ctc ttc atg cta aca gag gcc ctg ctc tcc cgg ggt ctg gtg gtg 2592
Asn Leu Phe Met Leu Thr Glu Ala Leu Leu Ser Arg Gly Leu Val Val
850 855 860

tcc ggc cac gac gtc agc gac gga ggc atg gtg gtt acc gct atc gag 2640
Ser Gly His Asp Val Ser Asp Gly Gly Met Val Val Thr Ala Ile Glu
865 870 875 880

atg gcc ctg gcc ggt aac cga ggc cta cag att cgc atc ccg tca gag 2688
Met Ala Leu Ala Gly Asn Arg Gly Leu Gln Ile Arg Ile Pro Ser Glu
885 890 895

gaa acg ccc ctg cag tgg ctg gtg tcc gaa acc ccg ggc gtc atc ttc 2736
Glu Thr Pro Leu Gln Trp Leu Val Ser Glu Thr Pro Gly Val Ile Phe
900 905 910

gaa atc cag ccc cag cac gtc gac gag gtg cgc cag gcg tgc caa aac 2784
Glu Ile Gln Pro Gln His Val Asp Glu Val Arg Gln Ala Cys Gln Asn
915 920 925

ttt gac tgc cgg gcc acc gtg tgc ggc acc gtg ggt cag gag ggt ctg 2832
Phe Asp Cys Arg Ala Thr Val Cys Gly Thr Val Gly Gln Glu Gly Leu
930 935 940

tca gaa aga atc gtc atc tca cac aac aac gat gaa gtc tat tca cag 2880
Ser Glu Arg Ile Val Ile Ser His Asn Asn Asp Glu Val Tyr Ser Gln
945 950 955 960

acc ctc act tcg gtg gcg gca aac tgg acc tcg ttc tcg gac gaa cag 2928
Thr Leu Thr Ser Val Ala Ala Asn Trp Thr Ser Phe Ser Asp Glu Gln
965 970 975

tgg tac tca tgg ggg ccc agc ttc acg ccc gcg cag gaa ctc tac aga 2976
Trp Tyr Ser Trp Gly Pro Ser Phe Thr Pro Ala Gln Glu Leu Tyr Arg
980 985 990

aag gac tac ggg tgc aac caa cac aac ctg ggc cac ctc gcc gag gtt 3024
Lys Asp Tyr Gly Cys Asn Gln His Asn Leu Gly His Leu Ala Glu Val
995 1000 1005

tgc aga aac agc gaa ctg act ctg ttt gca acg ccg tcg agg ccg cca 3072
Cys Arg Asn Ser Glu Leu Thr Leu Phe Ala Thr Pro Ser Arg Pro Pro
1010 1015 1020

gca gtg gcc gcc ttg gtc acc ccc ggt gcc ccg ctt ccc agg gcg ctg 3120
Ala Val Ala Ala Leu Val Thr Pro Gly Ala Pro Leu Pro Arg Ala Leu
1025 1030 1035 1040

atg gcc gcc ttc acc aac gtc ggg ttc gac gtc gcg gcc gtc tcc acc 3168
Met Ala Ala Phe Thr Asn Val Gly Phe Asp Val Ala Ala Val Ser Thr
1045 1050 1055

gac gac ctt aga gga ggg aac atc ctc cgg gga ttc tcc ggt ctg acc 3216
Asp Asp Leu Arg Gly Gly Asn Ile Leu Arg Gly Phe Ser Gly Leu Thr
1060 1065 1070

atc ggt gga acc gtg ggc atc gaa gac agc tac gtg ggg gcc cga tgc 3264
Ile Gly Gly Thr Val Gly Ile Glu Asp Ser Tyr Val Gly Ala Arg Cys
1075 1080 1085

gcg atc atg ggt ctt ctc aac gac cct ggg tgc tat ggg ggc ctt atg 3312
Ala Ile Met Gly Leu Leu Asn Asp Pro Gly Cys Tyr Gly Gly Leu Met
1090 1095 1100

gcc ttc ttc cgc cga gcg gat aca ttt tca cta tgc tgc ggg gag ttt 3360
Ala Phe Phe Arg Arg Ala Asp Thr Phe Ser Leu Cys Cys Gly Glu Phe
1105 1110 1115 1120

ggg ttc cag ctc ctg ggg gcg ctc ggt ctt ctt aga gag acc cct cac 3408
Gly Phe Gln Leu Leu Gly Ala Leu Gly Leu Leu Arg Glu Thr Pro His
1125 1130 1135

gac acg ccc gga ccc aaa acc ccc gac cag tgg gac atc cac ctg gag 3456
Asp Thr Pro Gly Pro Lys Thr Pro Asp Gln Trp Asp Ile His Leu Glu
1140 1145 1150

gag aac gca tca ggg aac cac gag tgc ctc tgg cta aac ctc cac atc 3504
Glu Asn Ala Ser Gly Asn His Glu Cys Leu Trp Leu Asn Leu His Ile
1155 1160 1165

ccc cag acc acc atc agc atc atg ttt agg gtt ctt cgg ggc ctc gtc 3552
Pro Gln Thr Thr Ile Ser Ile Met Phe Arg Val Leu Arg Gly Leu Val
1170 1175 1180

ctc cca ggg tgg gcc aac ggc cgg tac ctg ggc gtg aga tac ccg cgc 3600
Leu Pro Gly Trp Ala Asn Gly Arg Tyr Leu Gly Val Arg Tyr Pro Arg
1185 1190 1195 1200

gac gcc atg gag tac cat ctc aac cag cag cag cgc atc gcc ctg aac 3648
Asp Ala Met Glu Tyr His Leu Asn Gln Gln Gln Arg Ile Ala Leu Asn
1205 1210 1215

tac cac acg ggc aac gcc gac ccg agg atg ttc gcg cag cac tac cca 3696
Tyr His Thr Gly Asn Ala Asp Pro Arg Met Phe Ala Gln His Tyr Pro
1220 1225 1230

cgc aac ccg tct gcc aac tcg gcc gtg gcc gcc atc acg tca ccg gac 3744
Arg Asn Pro Ser Ala Asn Ser Ala Val Ala Ala Ile Thr Ser Pro Asp
1235 1240 1245

ggg cgc cac ctg gcc tcc ctg gtc gac ccc gcc gtg acg ttc cac ccc 3792
Gly Arg His Leu Ala Ser Leu Val Asp Pro Ala Val Thr Phe His Pro
1250 1255 1260

tgg cag tgg gcg tac gtg ccg ccc gag ctt gcg gac atg aca gtt tcc 3840
Trp Gln Trp Ala Tyr Val Pro Pro Glu Leu Ala Asp Met Thr Val Ser
1265 1270 1275 1280

ccg tgg gcg ctg gcc ttc cag tcc ctg ttc ctc tgg tgc ata aga aac 3888
Pro Trp Ala Leu Ala Phe Gln Ser Leu Phe Leu Trp Cys Ile Arg Asn
1285 1290 1295

aga cag taa 3897
Arg Gln

<210> 165
<211> 1298
<212> PRT
<213> Macaca mulatta rhadinovirus 17577

<400> 165
Met Ala Gln Arg Thr Asn Pro Arg Trp Ala Ala Ala Ala Leu Ser Pro
1 5 10 15
Glu Glu Glu Ala Phe Ile His Asp Asn Ser Asp Ala Glu Ser Val Leu
20 25 30
Ala Leu Val Pro Glu Gln Cys Phe Ser Glu Phe Leu Leu Trp Leu Val
35 40 45
Thr Arg Pro Ser Asp Asn Phe Asp Asn Asp Asp Asp Asp Pro Ala Leu
50 55 60
Gly Val Ile Trp His Leu Leu Ala Pro Leu Val Asn Tyr Ala Pro Leu
65 70 75 80
Glu Thr Arg Ser Ala His Leu Gln Gly His His Thr Ile Ser Leu Pro
85 90 95
Tyr Gly Pro Asp Leu Met Arg Gln Pro Thr Thr Arg Ser Ser Glu Ile
100 105 110
Val Gln Cys Leu Arg Asp Ser Gly Leu Asp Arg Thr Leu Arg Leu Glu
115 120 125
Val Gly Arg His Leu Ser Cys Gln Thr Arg Arg Phe Val Ala Asp Arg
130 135 140
Val Pro Pro Gly Thr Leu Ala Ala Leu Thr Leu Gly Thr Leu Val Glu
145 150 155 160
Tyr Asp Val Arg Val Gln Arg Gln Leu Pro Val Thr Leu Gln Ser Thr
165 170 175
Ala Trp Arg Pro Leu Pro Glu Arg Asp Pro Ile Cys Ala Ala Val Met
180 185 190
Leu Pro Leu Gln Arg Asn Ile Leu Pro Leu Ala Val Gln Ala Ser Asn
195 200 205
Gly Asn Ser Tyr Thr Val Ser Arg Tyr Ala Val Met Ala Arg Arg Ser
210 215 220
Tyr Ser Cys Val Phe Gln Arg Leu Pro Cys Glu Asn Val Thr His Ile
225 230 235 240
Ala Asp Ser Phe Thr His Leu His Ser Ala Ile Gln Thr Gly Ala Gly
245 250 255
Ala Leu Gln Asn Ile Leu Phe His Ala Thr Leu Leu Pro Gly Gly Glu
260 265 270
Ile Arg Ser Ala Leu Cys Gly Phe Tyr Ala Thr Thr Pro Ser Val Gly
275 280 285
Ala Phe Ser Arg Ala Arg His Arg Ala Ile Asn Thr Thr Ala Thr Leu
290 295 300
His Cys Gln Gln Leu Ala Arg Thr Gly Thr Pro Val Leu Gly Gly Phe

305 310 315 320
Leu Lys Thr Val His Ser Ala Thr Thr Ser Glu Ala Asn Val Ile Thr
325 330 335
Thr Thr Ser Leu Leu Ser Cys Val Pro Gln Ala Tyr Thr Phe Leu Arg
340 345 350
Arg Ser Leu Phe Ser Gln Pro Ile Ile Cys Leu Gly Ser Phe Glu Pro
355 360 365
Val Asp Gly Asp Gly Asn Gln Arg Ser Leu Tyr Leu Gly Ser Ala Ala
370 375 380
Gly Ile Thr Arg Ile Thr Gln Thr Leu Ser Leu Ala Tyr Glu Ile Leu
385 390 395 400
Glu Gly Pro Leu Phe Thr Ser Ile Asn Arg Ala His Glu Pro Ala Ser
405 410 415
Val Ile Gly His Leu Gly Ala Leu Val Ser Arg Gly Gly Leu Arg Leu
420 425 430
Phe Val Ser Gln Leu Pro Pro Thr Ile Leu Ser Gln Leu Thr Ala Thr
435 440 445
Pro Asp Ile Ser Arg Glu Thr Val Asn Asp Ile Leu Val Asn Lys Phe
450 455 460
Leu Asn Val Ser Ala Cys Val Val Phe Ala Val Leu Pro Arg Asp Thr
465 470 475 480
Glu Pro Glu Pro Gly Pro Leu Asp Ala Ile Arg Arg Ala Ala Arg Ile
485 490 495
Cys Gly Cys Pro Phe Ala Val Val Gly Glu Thr Cys Glu Glu Leu Gly
500 505 510
Ile Gln Phe Val Asn Asp Leu Glu Leu Trp Asn Pro Gly Ala Trp Pro
515 520 525
Ile Arg Gln Pro Thr Ser Ala Glu Val Ile Ala Thr Phe Gly Phe Asp
530 535 540
Glu Gln Pro Val Ser Ser Asn Trp Leu Val Arg Pro Glu Glu Pro Glu
545 550 555 560
Asp Gly Gly Glu Gln Ala Pro Ser Pro Thr Asp Trp Gly Leu Phe Arg
565 570 575
Leu Ala Ser Val Val Asp Gln Leu Leu Arg Cys Pro Thr Val Gly Ser
580 585 590
Lys Glu Phe Val Thr Arg His Val Asp Arg Cys Ser Asn Gly Leu Val
595 600 605
Ala Gln Gln Cys Glu Val Gly Pro Leu Gly Arg Pro Leu Ser Asp Tyr
610 615 620
His Ile Val Asn His Thr Ser Val Phe Thr Asp Arg Met Ala Arg Val
625 630 635 640
Pro Ile Tyr Arg Pro Gln Pro Ile Thr Arg Gln Asp Ala Thr Glu Arg
645 650 655
Leu Val Ser Pro Glu Thr Trp Val Thr Gln Gly Arg Gly Arg Asn Arg
660 665 670
Trp Val Gly Gln Cys Val Ala Tyr Gly Glu Gln Ala Tyr Lys Met Gly
675 680 685
Ile Asn Ala Ala Val Gly Ala Arg Tyr Ala Ile Cys Glu Ala Val Thr
690 695 700
Asn Ile Met Leu Ala His Val Arg Arg Leu Ser Asp Ile Thr Leu Thr
705 710 715 720
Ala Ser Val Gly Trp Asn Pro Glu Asp Asp Gln Ala Trp Leu Leu Gln
725 730 735
His Thr Leu Phe Ala Cys Lys Glu Leu Cys Arg Asp Leu Ser Ile Asn
740 745 750
Phe Ala Ile Thr Ser Ala Gly Ser Thr Pro Cys Leu Ser Glu Glu Leu
755 760 765
Ile Ser Ala Thr Gln Gln His Gln Thr Val Ala Pro Val Pro Phe Asn
770 775 780
Ala Val Val Ile Thr Ala Thr Ala Glu Val Lys Ser Ser Arg Gln Arg
785 790 795 800

Val Thr Pro Asp Leu Lys Ala Thr Gly Asn Leu Ile Val Leu Val Ser
805 810 815
Phe Pro Val Pro His Leu Thr Gln Gly Ser Thr Phe Glu His Leu Cys
820 825 830
Leu Leu Pro Ser Pro Thr Leu Pro Asp Val Gln Ala Thr His Leu Ala
835 840 845
Asn Leu Phe Met Leu Thr Glu Ala Leu Leu Ser Arg Gly Leu Val Val
850 855 860
Ser Gly His Asp Val Ser Asp Gly Gly Met Val Val Thr Ala Ile Glu
865 870 875 880
Met Ala Leu Ala Gly Asn Arg Gly Leu Gln Ile Arg Ile Pro Ser Glu
885 890 895
Glu Thr Pro Leu Gln Trp Leu Val Ser Glu Thr Pro Gly Val Ile Phe
900 905 910
Glu Ile Gln Pro Gln His Val Asp Glu Val Arg Gln Ala Cys Gln Asn
915 920 925
Phe Asp Cys Arg Ala Thr Val Cys Gly Thr Val Gly Gln Glu Gly Leu
930 935 940
Ser Glu Arg Ile Val Ile Ser His Asn Asn Asp Glu Val Tyr Ser Gln
945 950 955 960
Thr Leu Thr Ser Val Ala Ala Asn Trp Thr Ser Phe Ser Asp Glu Gln
965 970 975
Trp Tyr Ser Trp Gly Pro Ser Phe Thr Pro Ala Gln Glu Leu Tyr Arg
980 985 990
Lys Asp Tyr Gly Cys Asn Gln His Asn Leu Gly His Leu Ala Glu Val
995 1000 1005
Cys Arg Asn Ser Glu Leu Thr Leu Phe Ala Thr Pro Ser Arg Pro Pro
1010 1015 1020
Ala Val Ala Ala Leu Val Thr Pro Gly Ala Pro Leu Pro Arg Ala Leu
1025 1030 1035 1040
Met Ala Ala Phe Thr Asn Val Gly Phe Asp Val Ala Ala Val Ser Thr
1045 1050 1055
Asp Asp Leu Arg Gly Gly Asn Ile Leu Arg Gly Phe Ser Gly Leu Thr
1060 1065 1070
Ile Gly Gly Thr Val Gly Ile Glu Asp Ser Tyr Val Gly Ala Arg Cys
1075 1080 1085
Ala Ile Met Gly Leu Leu Asn Asp Pro Gly Cys Tyr Gly Gly Leu Met
1090 1095 1100
Ala Phe Phe Arg Arg Ala Asp Thr Phe Ser Leu Cys Cys Gly Glu Phe
1105 1110 1115 1120
Gly Phe Gln Leu Leu Gly Ala Leu Gly Leu Leu Arg Glu Thr Pro His
1125 1130 1135
Asp Thr Pro Gly Pro Lys Thr Pro Asp Gln Trp Asp Ile His Leu Glu
1140 1145 1150
Glu Asn Ala Ser Gly Asn His Glu Cys Leu Trp Leu Asn Leu His Ile
1155 1160 1165
Pro Gln Thr Thr Ile Ser Ile Met Phe Arg Val Leu Arg Gly Leu Val
1170 1175 1180
Leu Pro Gly Trp Ala Asn Gly Arg Tyr Leu Gly Val Arg Tyr Pro Arg
1185 1190 1195 1200
Asp Ala Met Glu Tyr His Leu Asn Gln Gln Gln Arg Ile Ala Leu Asn
1205 1210 1215
Tyr His Thr Gly Asn Ala Asp Pro Arg Met Phe Ala Gln His Tyr Pro
1220 1225 1230
Arg Asn Pro Ser Ala Asn Ser Ala Val Ala Ala Ile Thr Ser Pro Asp
1235 1240 1245
Gly Arg His Leu Ala Ser Leu Val Asp Pro Ala Val Thr Phe His Pro
1250 1255 1260
Trp Gln Trp Ala Tyr Val Pro Pro Glu Leu Ala Asp Met Thr Val Ser
1265 1270 1275 1280
Pro Trp Ala Leu Ala Phe Gln Ser Leu Phe Leu Trp Cys Ile Arg Asn

1285 1290 1295
Arg Gln

<210> 166
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 166
cctatgggct ccatgagc 18

<210> 167
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 167
atcgtcaatc aggctgcg 18

<210> 168
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 168
atattaaaca ctcgccgc 18

<210> 169
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 169
atgaggggcc ttttcgtgtg c 21

<210> 170
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 170
ctgaatcccg ctgccaaggc c 21

<210> 171
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 171
atgttccctg tctggttcgt c 21

<210> 172
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 172
ttacatcata gctattgcgc g 21

<210> 173
<211> 271
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 173
gcggcgcccg tgtttttggc ctgctgccag gtggcagccg tccaggcgcg tcactaccga 60
aacccgtgct ctgcggccac tatcttcgcc agcgccacca ccggcctcct ctgcggcgcg 120
cgtcaccgaa accccaaaca tgatgcttta ggttgcgttt ggggatatct ttcggtgccc 180
tccgaacgac ccggcgcggc ggggattttt ccgggtcgtt cggagggcac cgaaagatgt 240
tgccacgcgc gcgactggcg accgtgaaaa c 271

<210> 174
<211> 304
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 174
gcctagctcc taatgtttgc cttgccgcct agctcctaat gtttgccttg ccgcctagct 60
cctaatgttt gccttgccgc ctagctccta atgtttgcct tgccgcctag ctcctaatgt 120
ttgccttgcc gcctagctcc taatgtttgc cttgccgcct agctcctaat gtttgccttg 180
ccgcctagct cctaatgttt gccttgccgc ctagctccta atgtttgcct tgccgcctag 240
ctcctaatgt ttgccttgcc gcctagctcc taatgtttgc cttgccgcct agctcctaat 300
gttt 304

<210> 175
<211> 1008
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 175

gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag 60
ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc 120
tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc 180
ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc 240
cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc 300
gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag 360
ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc 420
tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc 480
ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc 540
cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc 600
gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag 660
ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc 720
tccccgttcc ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc 780
ccgagggtcc cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc 840
cgggctcccc gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc 900
gttccccgag ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccgag 960
ggtcccgggc tccccgttcc ccgagggtcc cgggctcccc gttccccg 1008

<210> 176
<211> 405
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 176
ggtggggtgc aagtcccccg gtggggtgca agtcccccgg tggggtgcag gtcccccggt 60
ggggtgcagg tcccccggtg gggtgcaagt cccccggtgg ggtgcaagtc ccccggtggg 120
gtgcaagtcc cccggtgggg tgcaagtccc ccggtggggt gcaagtcccc cggtggggtg 180
caagtccccc ggtggggtgc aagtcccccg gtggggtgca agtcccccgg tggggtgcac 240
gtcccccggt ggggtgcaag tcccccggtg gggtgcaagt cccccggtgg ggtgcaagtc 300
ccccggtggg gtgcaagtcc cccggtgggg ttcaagtccc ccggtggggt gcaagtcccc 360
cggtggggtg caagtccccc ggtggggtgc aggtcccccg gtggg 405

<210> 177
<211> 1029
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 177
gcggctcggc tccggggtgg ctccgggtgg gggcggctcg gctccggggt ggctccgggt 60
gggggcggct cggctccggg gtggctccgg gtgggggcgg ctcggctccg ggtggctcc 120
gggtgggggc ggctcggctc cggggtggct ccgggtgggg gcggctcggc tccggggtgg 180
ctccgggtgg gggcggctcg gctccggggt ggctccgggt gggggcggct cggctccggg 240
gtggctccgg gtgggggcgg ctcggctccg ggtggctcc gggtgggggc ggctcggctc 300
cggggtggct ccgggtgggg gcggctcggc tccggggtgg ctccgggtgg gggcggctcg 360
gctccggggt ggctccgggt gggggcggct cggctccggg gtggctccgg gtgggggcgg 420
ctcggctccg ggtggctcc gggtgggggc ggctcggctc cggggtggct ccgggtgggg 480
gcggctcggc tccggggtgg ctccgggtgg gggcggctcg gctccggggt ggctccgggt 540
gggggcggct cggctccggg gtggctccgg gtgggggcgg ctcggctccg ggtggctcc 600
gggtgggggc ggctcggctc cggggtggct ccgggtgggg gcggctcggc tccggggtgg 660
ctccgggtgg gggcggctcg gctccggggt ggctccgggt gggggcggct cggctccggg 720
gtggctccgg gtgggggcgg ctcggctccg ggtggctcc gggtgggggc ggctcggctc 780
cggggtggct ccgggtgggg gcggctcggc tccggggtgg ctccgggtgg gggcggctcg 840
gctccggggt ggctccgggt gggggcggct cggctccggg gtggctccgg gtgggggcgg 900
ctcggctccg ggtggctcc gggtgggggc ggctcggctc cggggtggct ccgggtgggg 960
gcggctcggc tccggggtgg ctccgggtgg gggcggctcg gctccggggt ggctccgggt 1020
gggggcggc 1029

<210> 178

<211> 196
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 178
ggtaagggct aggggtaagg gcaaggggta agggctaggg gtaagggcaa ggggtaaggg 60
caaggggtaa gggcaagggg taagggcaag ggtaagggc aaggggtaag ggcaaggggt 120
aagggcaagg ggtaagggca aggggtaagg gctaggggta agggctaggg gtaagggcta 180
ggggtaaggg ctaggg 196

<210> 179
<211> 1387
<212> DNA
<213> Macaca mulatta rhadinovirus 17577

<400> 179
cctagcggcc gggcccggag gcgcccggcc ggcgccgccg ccgggccgcg gccgccatct 60
tgcgcccggg gcgagggtcc cccgcgcgcc ccccgggccc gcgcgccggg cgccgccggc 120
ctcccccggc tcccggccct ccgccccgcc ggcccgggcc cgcggccgcc gcccccggcc 180
gccccgcgcc gcgcgccggg ccgctttcgg ttcgcggggc cggggtccc gcggggggcc 240
ggggccgccc ccgggtgccg ctccgccggg cccggccgac tcccgggagc gccccggtcc 300
ggccgagccc ggcggcgccg cgcccccgacg ccccccgggc ccggggcccc acaagccgcg 360
gcgcgcggct cccgatgccg ggcggccgcc gcccggcatg gcggtcctcc gccggcctcc 420
cctcccccac gccgccccga aaggtggtct ccgcgccgcc gggagggggg ccggggcccg 480
gggcgcctcg gcggggcccg gcgcggggcg cgaccgaggg ccccgggaga acgggggatc 540
gggaaaacgc gaggggagcg ggggacaggg gacggcgtgt gcgtgcttgt gagacaccgg 600
gtacggctgc ctgcctgctc gctggcctgc ttgctgaggg gacagtaggc ctgcttgctc 660
gctggcctgc ttgctgaggg gacagtaggc ctgcttgctg aggggacagt aggcctgctt 720
gctcgctggc ctgcttgctg aggggacagt agggctgctg gcttgctagt agggctgctc 780
gctggcctgc ttgcttgctc gctggcctgc ttgcttgctc gctggcctgc ttgcttgctc 840
gctggcctgc ttgcttgctc gctggcctgc ttgcttgctc gctggcctgc ttgcttgctc 900
gctggcctgc ttgcttgctc gctggcctgc ttgctgaggg gacagtaggg ctgcttgctt 960
gctaagggga cggtacgcct gcctgatggc ttgatagtag ggctgctggg ctgctagtag 1020
ggctgctggg ctgctagtag ggctgctggg ctgctagtag ggctgctggg ctgctagtag 1080
ggctcctggg ctgctagtag ggctgctggg ctgctagtag ggctcctggg ctgctagtag 1140
ggctgctggg ctgctagtag ggctgctggg ctgctagtag ggctgctggg ctgctagtag 1200
ggctgcctgc tggcttgctt gcttgcttgc tagtggggcc gcttgcctgc tactagggct 1260
gctgtgcagc tgggagaaca gagtagggct gccggccagc tgcgtgcgag ggcgtccgag 1320
ggccagacga ggacacggga cccgggcctc tcccccgccc ggaccgccgg gcacccggcc 1380
cagatct 1387

09831000.050201

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)___________ of ______________________________ were not present for scanning.
(Document title)

☐ Page(s)___________ of ______________________________ were not present for scanning.
(Document title)

☒ ***Scanned copy is best available.*** SOME DRAWINGS ARE DARK